chinaXiv:202305.00174v1

基于规则的知识图谱推理研究综述

曾泽凡[1,2] 成清[1] 司悦航[1] 刘忠[1]
[1](国防科技大学大数据与决策实验室 长沙 410073)
[2](福建省军区 福州 350001)

摘要：
［**目的**］当前，基于规则的知识图谱推理方法不断涌现，但缺乏系统的归类和分析。本文梳理了基于规则的知识图谱推理相关研究工作。
［**方法**］对相关经典方法的核心思路、关键技术和优缺点进行阐述分析，并对不同类型方法的性能进行比较。
［**结果**］本文归纳出 4 类方法，包括基于归纳逻辑编程的方法、基于概率图和规则的方法、基于嵌入表示和规则的方法和基于神经网络和规则的方法。
［**局限**］相关方法的具体应用研究目前还比较少。
［**结论**］基于规则的知识图谱推理方法通过挖掘底层逻辑规则执行推理，泛化能力和可解释性较强，基于规则的混合推理方法在未来仍有较大应用前景。
关键词：知识图谱推理 规则 归纳逻辑编程 概率图 嵌入表示 神经网络
分类号：TP391

A Survey on Research of Rule-Based Knowledge Graph Reasoning

ZENG Ze-fan[1,2] CHENG Qing[1] SI Yue-hang[1] LIU Zhong[1]
[1](Big Data and Decision Laboratory, National University of Defense Technology, Changsha 410073)
[2](Fujian Province Military Region, Fuzhou 350001)

Abstract:
[Objective] Currently, rule-based knowledge graph reasoning methods are constantly emerging, but there is a lack of systematic classification and analysis. This article summarizes the research work related to rule-based knowledge graph inference.
[Methods] Elaborate and analyze the core ideas, key technologies, advantages and disadvantages of relevant classic methods, and compare the performance of different types of methods.
[Results] This paper summarized four kinds of methods, including methods based on inductive logic programming, methods based on probability graphs and rules, methods based on embedding representation and rules, and methods based on neural networks and rules.
There is currently relatively little research on the specific application of relevant methods.
[Conclusions] The rule-based knowledge graph reasoning method performs inference by mining underlying logical rules, with strong generalization ability and interpretability. The rule-based hybrid reasoning method still has great application prospects in the future.
Keywords: knowledge graph reasoning, rules, inductive logic programming, probability graph, embedding representation, neural network

1 引言

知识表示和学习是人类认识世界、学习技能和做出决策的必要手段。大数据时代催生了蕴藏着丰富知识的海量数据，如何利用好这些知识辅助人类解决问题，

成为了大数据背景下的一大难题。2012 年，谷歌首次提出了知识图谱(Knowledge Graph, KG)的概念，它是一种简单而有效的知识表示方法，包含了丰富的人类知识资源[1]。近年来，互联网技术和 Web 应用迅速发展，大量数据被发布到互联网上，成为大规模知识提取和知识图谱构建的重要来源。随着 WordNet[2]、DBpedia[3]、YAGO[4]、NELL[5]、Probase[6]和 CN-DBpedia[7]、zhishi.me[8]等国内外开放知识库迅猛增长，知识图谱已被广泛用于支持语义网络的实际应用，如推荐系统[9]、问答系统[10]、智能对话[11]等。

根据实体类型和要素的不同，知识图谱可以分为静态知识图谱(Static Knowledge Graph, SKG)、时序知识图谱(Temporal Knowledge Graph, TKG)和多模态知识图谱(Multi-Modal Knowledge Graph, MMKG)三类[12]，其中最常见且研究工作最为密集的是静态知识图谱。大多数研究中，静态知识图谱在不与其他两类知识图谱同时提及时，通常简称为知识图谱。本文的研究对象为静态知识图谱，为方便表示，本文之后出现所有的“知识图谱”均指代“静态知识图谱”。

受限于知识来源的不全面性和知识抽取过程中的偏差，当前大多数知识图谱都不同程度存在缺失或错误问题，因此知识图谱推理(Knowledge Graph Reasoning, KGR)技术[13]应运而生。知识图谱推理是基于知识图谱从已有事实中推导出新的事实的推理方法，广泛运用于智能问答、推荐系统、对话生成、信息提取和图像分类[14-18]等人工智能任务中，在军事、金融、交通运输和网络安全等领域[19-22]都发挥着至关重要的作用。

为实现从“计算智能”、“感知智能”到“认知智能”的跃变[23]，可解释人工智能(eXplainable Artificial Intelligence, XAI)在科研学术界受到的关注和重视程度不断增加，被视作人工智能继续稳步发展的必要条件。尽管目前研究较多的基于分布式表示的知识图谱推理方法[24,25]效率较高，但其在潜在空间中操作和模型的“黑盒”特性，也为直观理解推理依据和决策过程带来了很大的挑战。因此，此类方法在一些对可靠性要求较高的领域(如军事、医疗和信息安全等)难以适用。

基于规则的推理方法直接通过挖掘底层逻辑规则推导出新的事实，在保证推理准确性的同时，还具有较强的可解释性，能使用户更直观地理解推理过程，为辅助决策提供可靠指导。此外，基于规则的知识推理还能实现未出现过事实的推理(即归纳式推理)。早期基于规则的知识图谱推理依赖于先验知识、硬匹配和遍历搜索，导致规则的表达能力有限，扩展性和鲁棒性较差，效率也比较低[26]。随着深度学习、强化学习[27,28]等技术方法的日臻优越，越来越多的研究通过结合这些技术进行规则学习或利用规则辅助相关推理方法，以实现高效、准确、可迁移、可解释的知识图谱推理。

ChatGPT 领衔的大语言模型(Large Language Model, LLM)为知识推理提供了新的工具[29-31]。LLM 能够学习海量的知识，利用预训练语料库中丰富的事实知识或外部数据检索实现高效的知识图谱推理。但 LLM 也存在受到“幻觉”(hallucination)[32]即错误、混淆的信息影响的潜在风险。学习可信的规则并将其作为“指令”融入大模型中，开发基于 LLM 的神经-规则混合推理模型，以辅助 LLM 完成知识推理，有望进一步提升大模型的可靠性、可理解性和领域特定适用性。

目前，学者们已总结研究并发表了多篇知识图谱推理相关的综述文章。其中，文献[12]详实地梳理归纳了静态、时序和多模态知识图谱推理的研究工作；文献[33]区分多样本、小样本和单(零)样本对静态知识图谱推理进行综述；文献[34]对基于神经网络的知识图谱推理方法进行系统介绍；文献[35]搜集总结了图神经

chinaXiv:202305.00174v1

网络应用于知识图谱推理的方法和关键技术；文献[36]聚焦基于表示学习的知识图谱推理，归纳比较了相关方法；文献[23]和[26]从可解释的角度，梳理并分类对比了不同的知识图谱推理技术及其解释方法；文献[37]对基于深度强化学习的知识图谱推理研究进展进行综述；文献[38]分类介绍了基于符号和统计的混合知识图谱推理方法；文献[39]综述了如何结合逻辑规则与嵌入技术进行知识图谱推理；文献[40]则从符号主义和连接主义两种范式的角度，综述了符号推理、神经推理和神经符号混合推理相关方法。

本文以基于规则的知识图谱推理方法为主要调研对象，归纳整理了与基于规则的知识图谱推理相关的文献，在介绍背景知识和基本概念的基础上，详细阐述基于归纳逻辑编程的知识图谱推理、基于概率图模型和规则的知识图谱推理、基于嵌入表示和规则的知识图谱推理以及基于神经网络和规则的知识图谱推理方法，方法分类框架如图 1 所示。重点介绍各类方法中典型工作的核心思路、关键技术和推理效果，并对不同类型方法进行了比较分析，总结展望主要挑战和未来的研究方向，以期为下一步的研究工作提供参考借鉴。

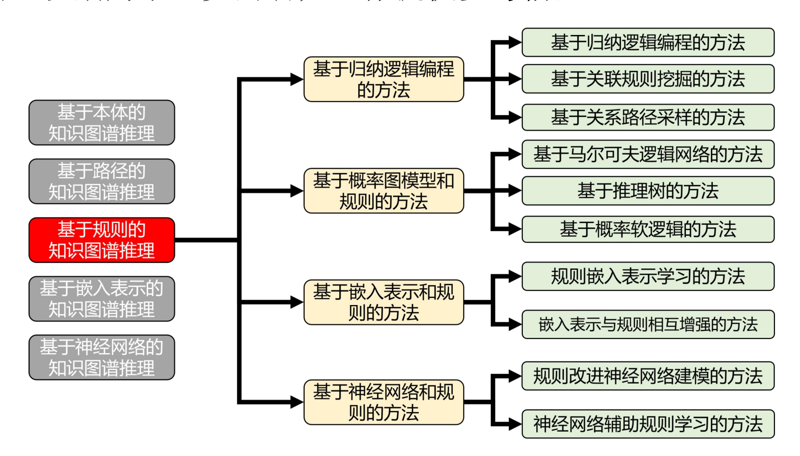


图 1　基于规则的知识图谱推理方法分类框架

2　背景知识和概念定义

本节主要介绍知识图谱、知识图谱推理和基于规则的知识推理相关的背景知识和基本概念。

2.1　知识图谱

知识图谱是一种图形化的知识表示方式，是由节点和边组成巨型语义网，其中节点表示物理世界中的概念或实体，边表示节点之间的拓扑连接和语义关系。文献[12]将知识图谱分为静态知识图谱、时序知识图谱和多模态知识图谱三类。时序知识图谱是指包含时间戳信息的知识图谱，多模态知识图谱是指实体类型包含图片、文本等多模态数据的知识图谱。

知识图谱通常以三元组的形式表示事实，即(头实体，关系，尾实体)，其词汇表定义在模式(也称为本体)中，表明两个实体由特定的关系连接起来。例如，(Nunez,PlayforTeam,Liverpool)表示运动员努涅斯效力于利物浦足球队。

在符号表示上，知识图谱定义为$\mathcal{KG}=\{\mathcal{E},\mathcal{R},\mathcal{F}\}$，其中$\mathcal{E}$、$\mathcal{R}$和$\mathcal{F}$分别表示实体、关系和事实的集合。三元组$\langle e_h,r,e_t\rangle$表示一个事实$\mathcal{F}$，其中$e_h,e_t\in\mathcal{E},r\in\mathcal{R}$，$\mathcal{F}\subseteq\mathcal{E}\times\mathcal{R}\times\mathcal{E}$。在前述三元组中$e_h=Nunez,r=PlayforTeam,e_t=Liverpool$。此外，

chinaXiv:202305.00174v1

在知识图谱构建时，若 $e_h, r, e_t \in \mathcal{F}$，通常将 e_t, r^{-1}, e_h 也添加到 $\mathcal{F}$ 中。上标"-1"表示逆关系。

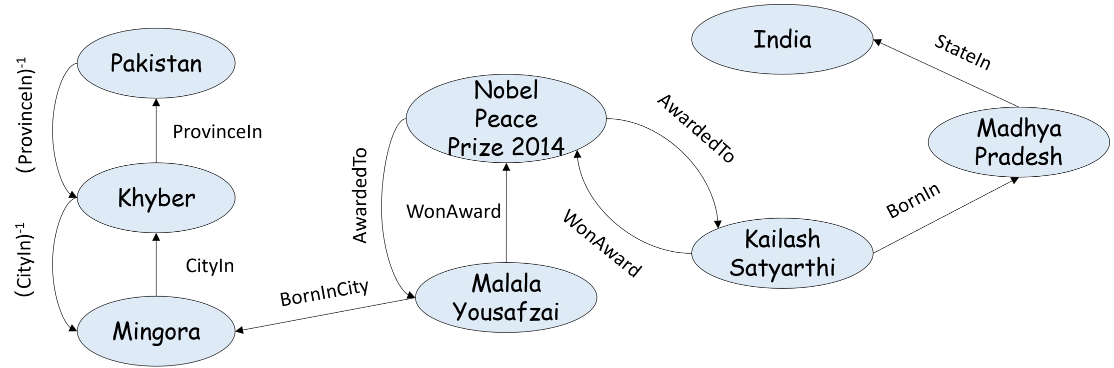


图 2　知识图谱示例

在图模型表示上，知识图谱可以视作一个有向标记多重图 $\mathcal{G}=\{V,E,\mathcal{R}\}$，$V$ 和 E 分别表示图的顶点和边，其中 $E \in V \times \mathcal{R} \times V$ 且 $V=\mathcal{E}$。在图 2 中给出了一个简单的知识图谱案例。

2.2　知识图谱推理

知识图谱推理是从现有的知识出发，运用一定的推理方法获取所蕴含的隐藏事实或新事实，其本质是通过推导或归纳将个体知识推广到一般性知识的过程[41]。根据对象三元组是否完整，可以将知识图谱推理任务分为知识图谱去噪和知识图谱补全两类，前者是对现有完整三元组正确与否进行判断，而后者是对缺失的实体或关系进行预测。

在符号表示上，给定一个知识图谱 $\mathcal{KG}=\{\mathcal{E},\mathcal{R},\mathcal{F}\}$，知识图谱推理就是利用已有的事实来推断查询的事实 $\langle e_h^q, r^q, e_t^q \rangle$。知识图谱去噪即判定 $\langle e_h^q, r^q, e_t^q \rangle$ 是否正确，知识图谱补全包括三类子任务，即头实体补全 $\langle ?, r^q, e_t^q \rangle$、尾实体补全 $\langle e_h^q, r^q, ? \rangle$ 和关系预测 $\langle e_h^q, ?, e_t^q \rangle$。图 3 给出了一个知识图谱推理的示例，红色箭头的关系相关的三元组对应去噪任务，黑色虚线箭头相关的三元组对应关系预测任务，蓝色箭头及其指向的实体对应需要补全的尾实体。

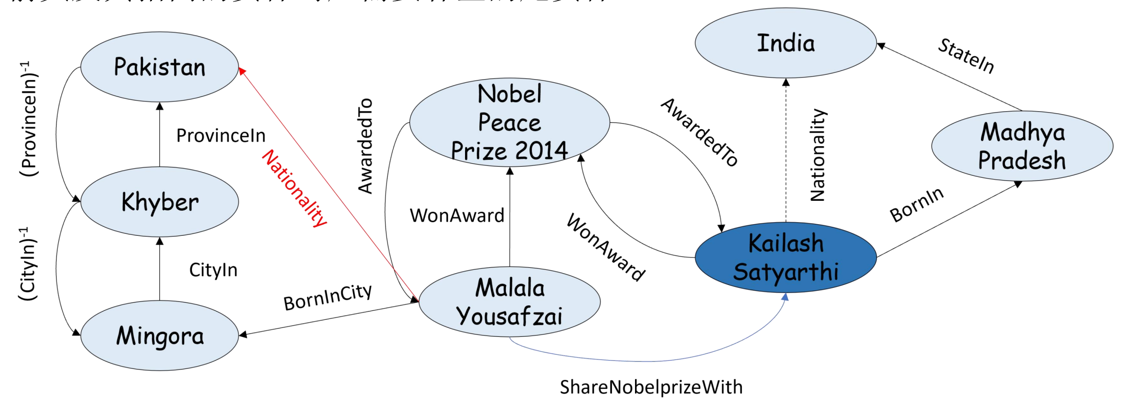


图 3　知识图谱推理示例

2.3　基于规则的知识图谱推理

基于规则的知识推理方法挖掘可靠的规则用于知识图谱推理。规则的基本单位称为原子(Atom)，通常用符号 α 表示。原子可以被定义为一元或多元(通常是二元)的谓词符号，将一系列变量映射为真(True)或假(False)。特别地，若一个原子中的所有变量都是常数项，该原子被称为基本原子(Grounding Atom)。多个原

chinaXiv:202305.00174v1

子利用逻辑连接词连接起来即构成一个规则。规则通常以 $body \Rightarrow head$ (规则体⇒规则头)的形式构成，规则体由多个原子的合取组成，表示规则成立的前件；规则头是包含目标谓词的单个原子，表示规则所得出的结论。

一阶逻辑(First Order Logic, FOL)[59]是知识表示和推理的一种主要形式，它将知识图谱中的三元组转化成"二元谓词+常量"的组合，即 <subject,predicate,object>≡predicate<subject,object>。例如在图 3 中，如果知道规则 $\text{BornInCity}\langle X,Y\rangle \wedge \text{CityIn}\langle Y,Z\rangle \wedge \text{ProvinceIn}\langle Z,W\rangle \rightarrow \text{Nationality}\langle X,W\rangle$，那么就可以依据此规则推理得出 Malala Yousafzai 的国籍为巴基斯坦。

霍恩规则逻辑(Horn Rule Logic)[58]也称为霍恩子句或霍恩规则，它是一阶逻辑的子集，具有形式简单、易于描述等特点，是一种广泛适用的规则表示方法。霍恩规则由头部和主体组成，其中头部是单个原子，主体是几个原子的连接。霍恩规则 H 可以表述为：

$$H : A\left(\alpha_1,...,\alpha_n\right) \rightarrow \alpha \,, \tag{1}$$

其中 α 称为头原子，$\alpha_1,...,\alpha_n (n>0)$ 为主体原子。A 为规则体，利用合取或析取等范式将主体原子连接在一起。规则体也称为公式，即若 A 为真，那么右边的规则头原子也为真。霍恩规则中的封闭规则指的是所有变量至少在两个原子中出现，这也是知识图谱推理中最常见的规则。

基于规则的知识推理是规则范式的泛化，在语义上符合自然语言表达和人类思维方式，可以提供清晰、明确的推理依据，具有较强的可解释性和较高的推理精度，是一种全局可解释的知识图谱推理方法。此类方法也是最早使用的知识图谱推理方法。然而，由于知识图谱语义的多样性，推理所涉及的规则通常较为复杂，因此传统的规则学习和知识推理方法也存在可扩展性差、表达能力有限、时间和空间复杂度高等问题。为解决上述问题，研究人员尝试着将图模型、神经网络等与规则相结合，利用其在可迁移性、鲁棒性和效率上的优势，进行混合推理。近几年来，混合推理已经成为知识图谱推理领域的研究热点之一。

3 基于归纳逻辑编程的知识图谱推理

归纳逻辑编程(Inductive Logic Programming, ILP)是基于一阶逻辑的规则学习方法[40]，这种方法旨在寻找数据中共享的逻辑程序、规则或公式所构成的底层模式。ILP 将逻辑规则约束为霍恩规则，然后根据给定的谓词、基础原子、正实例和负实例，构造能够解释正实例、拒绝负实例的规则。基于 ILP 的知识图谱推理方法可以分为基于一阶归纳学习的方法、基于关联规则挖掘的方法和基于关系路径采样的方法。

3.1 基于一阶归纳学习的方法

一阶归纳学习器(First Order Inductive Learner, FOIL)是一种领域独立的规则学习方法[42]，该方法旨在从文本中学习推理规则。FOIL 方法以 SHERLOCK 系统[43]为基础，按照如下 4 个步骤学习规则：(1)识别有效类及其实例；(2)发现类之间的关系；(3)利用类之间的关系学习推理规则并确定其置信度评分；(4)利用 HOLMES 推理引擎通过逻辑推理查找事实，并使用马尔可夫网络估计每个事实的置信度。FOIL 的具体流程步骤如图 4 所示。

chinaXiv:202305.00174v1

chinaXiv:202305.00174v1

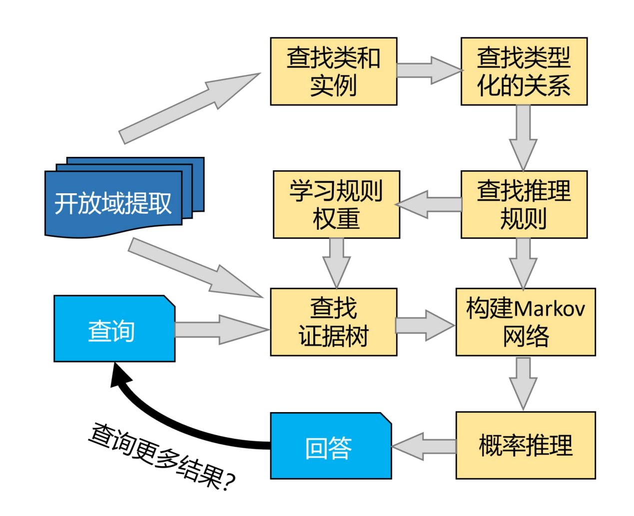


图 4　FOIL 方法的详细步骤

Landwehr 等[44]将朴素贝叶斯学习方法集成到 FOIL 中，提出了两种动态命题化规则学习方法——朴素贝叶斯一阶归纳学习器(Naive Bayesian First Order Inductive Learner, N-FOIL) 和树增强的朴素贝叶斯一阶归纳学习器(Tree-augmented Naive Bayesian First Order Inductive Learner, T-FOIL)。N-FOIL 不使用分治方法，计算复杂度与 FOIL 相同，但 N-FOIL 在每次迭代之后，实例的数量都会减少，并依据每次对额外子句(规则)的评分调整路径，有效增加了规则学习的多样性。为了防止模型参数过多，N-FOIL 中作出了“实例满足不同查询的概率彼此不相关”的强假设，即：

$$\mathbf{P}_\lambda(\mathbf{q}_1,\ldots,\mathbf{q}_k \mid \mathbf{p}) = \prod_i \mathbf{P}_\lambda(\mathbf{q}_i \mid \mathbf{p}), \tag{2}$$

$\mathbf{P}_\lambda$ 代表查询概率，$\mathbf{p}$ 为类标签，$\mathbf{q}_i$ 为查询，但该假设在实际情况中往往难以满足。T-FOIL 针对该问题，通过引入额外的依赖关系树来减少参数数量，即将式 2 松弛为：

$$\mathbf{P}_\lambda(\mathbf{q}_1,\ldots,\mathbf{q}_k \mid \mathbf{p}) = \prod_i \mathbf{P}_\lambda\left(\mathbf{q}_i \mid \mathbf{p}, \mathbf{q}_{\mathbf{pa}(i)}\right), \tag{3}$$

$\mathbf{pa}(i)$ 表示树增强朴素贝叶斯模型中节点 $\mathbf{q}_i$ 的附加父节点。树增强策略使得模型参数从 $\mathcal{O}\left(2^{\#nodes}\right)$ 降为 $\mathcal{O}(\#nodes)$，有效降低了计算复杂度，提升了学习的效率。

Landwehr 等[45]受核学习和函数学习之间关系的启发，将核函数学习类比为 ILP 统计关系学习中结构与参数学习的关系，进而把核学习技术与 FOIL 算法相结合，提出了 K-FOIL(Kernel First Order Inductive Learner)方法。在 K-FOIL 中，结构学习对应于在逻辑对象上推断一个合适的核，而参数学习对应于再现核 Hilbert 空间中的函数学习。然后，利用离散空间搜索算法在逻辑设置中诱导核函数 K，利用 K 执行多任务统计逻辑规则学习。实例分析表明，K-FOIL 在准确性和效率方面相对 FOIL 和 N-FOIL 都具有优势。

此外，Zeng 等[46]提出了 QuickFOIL 算法。该算法基于自顶向下贪婪 ILP 搜索，采用了驱动启发式搜索过程的评分函数和一个新的修剪策略，使算法能够找到高质量的规则。在构造子句时，从一个空文本开始，然后使用贪婪的启发式方法一次添加一个新的文本，然后修剪部分搜索空间的关键决策。QuickFOIL 将连接操作的数量从 $\mathcal{O}\left(n^k\right)$ 减少到 $\mathcal{O}(nk)$。此外，QuickFOIL 还开发了多种类查询优化和缓存技术，以加快查询的速度。

基于 FOIL 的方法在小规模的知识图谱中可以实现准确高效的推理，但在中大型知识图谱中，实体和关系的类型更加多样，规则的数量级也随之急剧增加，计算复杂度指数级上升，其准确度和效率也将大大下降。此外，FOIL 方法通常需要负实例，而知识图谱的开放世界假设(Open World Assumption, OWA)意味着缺失的数据不能用作负实例。

3.2 基于关联规则挖掘的方法

为解决 FOIL 方法存在的问题，提升推理方法的归纳效率，Galárraga 等[47]提出了不完全证据关联规则挖掘系统 AMIE(Association Mining under Incomplete Evidence)。AMIE 能够在大型 RDF(Resource Description Framework)知识库中挖掘 Horn 规则。AMIE 根据知识图谱关系的类型生成规则库，然后从知识图谱中找出满足该规则的实例，并计算规则的置信度，若置信度超过某个阈值，则认为规则可靠。

AMIE 规则的学习包括两个阶段，第一步是规则发现，将规则视为原子序列，初始化一个空的规则集，利用三种挖掘算子添加规则：(1)添加悬垂原子$\mathcal{O}_D$，即在已有的规则中添加新变量；(2)添加实例化原子$\mathcal{O}_I$，即在已有的规则中添加常量原子；(3)添加封闭原子$\mathcal{O}_C$，即添加一个新原子使规则不再扩展。第二步是规则修剪，根据部分完备性假设 PCA(Partial Completeness Assumption)计算置信度，若在一个规则中添加原子后置信度没有提升，那么规则不再扩展；若某个规则只有头部关系，那么判断其置信度是否超过阈值θ，若未超过θ则直接删除。AMIE 不需要知识图谱以外的知识，也不需要进行配置或参数调优，在运行时间、输出规则的数量和质量方面相比 FOIL 系列方法都有明显改进。

AMIE 系统通过约束规则长度、类型和置信度等指标以减小搜索空间，但由于在规则发现过程中所有关系都能通过规则与其他任意关系相联系，因此在面对大规模知识图谱时，仍面临着搜索空间爆炸的问题。为了弥补上述缺陷，Galárraga 等[48]人又提出了 AMIE 的改进版本 AMIE+。AMIE+主要在 3 个方面进行了优化：(1)规则发现步骤中，AMIE+在最后一步只添加$\mathcal{O}_C$和$\mathcal{O}_I$而不添加$\mathcal{O}_D$，以避免规则未封闭。(2)规则修剪步骤中，提出了一种置信度近似计算方法，用于计算复杂度较高规则的置信度。(3)将发现的规则与启发式类型检查和联合推理结合：类型检查在规则体中增加头尾实体所属类别信息，提高推理准确度；联合推理是指对于给定查询，如果多个规则都能推理得出候选答案，则通过聚合这些规则的 PCA 置信度以提高推理结果的排名。

AMIE+尽管准确度和效率相较 AMIE 有明显提升，但其算法执行过程中每次只能挖掘一条规则，无法实现并行化。针对该问题，Lajus 等[49]对其进行了进一步的改进，提出了 AMIE3 系统。该系统提出了多个复杂的剪枝策略、并行优化算法和置信度的惰性计算来加速策略挖掘，更适用于大规模知识图谱。此外，AMIE3 可以详尽地挖掘支持度和置信度阈值以上的所有规则，而无需求助于抽样或近似值。

此外，Wang 等[50]提出了一种规则学习方法 RDF2Rules，RDF2Rules 通过挖掘频繁谓词循环(Frequent Predicate Cycle, FPC)进行关联规则生成。该方法将定义了谓词循环(实体通过多个谓词回到原实体)$x_1, p_1d_1, x_2, ..., p_kd_k, x_1$，$x_i$表示实体，$p_i$表示谓词，$d_i \in \{-1,1\}$表示谓词方向，并通过支持度判定是否为 FPC。然后，利用 FPC 生成和评价规则，并在过程中使用实体类型信息，使得学习到的规则更加准确。

chinaXiv:202305.00174v1

然而，关联规则挖掘方法往往需要对所有的关系路径进行遍历，因此在应用于大规模知识图谱时计算复杂度高，效率低下。

3.3 基于关系路径采样的方法

为降低规则搜索的计算复杂度，学者们提出了基于关系路径采样的方法。这种方法的基本思想是，对每一个目标关系，只对知识图谱中与该关系相关联的三元组进行采样，然后利用采样信息进行规则学习。此类方法的代表性方法包括RuleN[51]、AnyBURL[52]及其变体和C-NN[55]等。

RuleN是一种基于双重随机采样的规则挖掘方法，支持两种类型的规则挖掘，为P_n类型规则和C类型规则，其形式为：

$$\begin{aligned} &P_n : r(x_1, x_{n+1}) \leftarrow s_1(x_1, x_2) \wedge \ldots \wedge s_n(x_n, x_{n+1}), \\ &C : r(x, a) \leftarrow \exists y\ r(x, y), \end{aligned} \tag{4}$$

该方法对于P_n类型待查询关系r，只抽取与之相关的K个三元组a,r,b，使用深度优先遍历算法来确定a和b之间长度为n的所有路径作为候选规则体。对满足P_n类型候选规则体的路径进行随机采样，判断路径中的头实体和尾实体之间是否存在关系r。对于C类型规则，同样随机选取K个三元组a,r,b，然后对于每个三元组创建规则$r(x,b) \leftarrow r(x,y)$和$r(a,y) \leftarrow r(x,y)$。最后，通过随机采样规则体为真实的三元组来计算路径规则的近似置信度。对于C类型规则，只需从r个事实中选取一个样本，并计算其主语和宾语位置中发现a或b的频率作为置信度。RuleN方法通过采样有效地减少了搜索空间，简化了计算过程。

Ferré提出了概念最近邻方法(Concepts of Nearest Neighbours, C-NN)方法。C-NN是一种基于实体聚类的规则学习方法，不需要大规模训练或重训练，避免了在运行时间和内存方面的高昂代价。C-NN引入了一种符号形式的K-近邻，用图形模式取代传统KNN的数字距离，图形模式提供了两个实体相似程度的可理解表示。对于每个实体的近邻筛选，本质上也是一种抽样过程。在规则学习中，基于公共图模式识别相似实体的集群，然后从这些图形模式中衍生出不同的规则，并结合起来进行推理。C-NN可以为每个推理提供解释。

Meilicke等提出了AnyBURL(Anytime Bottom-Up Rule Learning)方法。AnyBURL是一种能够随时从大规模知识图谱中学习逻辑规则的自底向上技术。方法的基本思想是从给定知识图谱$\mathcal{G}$中采样长度为n的路径，从$n=2$开始，从长度为n的路径中，学习长度为$n-1$的规则，路径上的第一条边对应于头部原子。达到长度为$n-1$的规则的一定饱和度SAT后，令$n=n+1$以学习更长的规则。其中，饱和度定义为$\mathrm{SAT}=|R'_S|/|R_S|$，$R_S$表示在当前时间跨度中发现的所有规则，$R'_S$表示当前时间跨度发现的与过去迭代中发现的规则重合的规则。AnyBURL的规则学习过程是在长度为t_s的时间跨度序列中进行的，在一个时间跨度内，算法通过迭代采样随机路径来学习尽可能多的规则。时间跨度结束后，对规则进行置信度评分。最后，通过生成所有候选规则的置信度的最大值对候选规则进行排序。该方法有较高的计算效率和较少的计算资源要求，并且能够根据提出的候选规则产生解释，显著优于其他基于规则的推理方法，在性能和可解释性方面有着巨大潜力。

AnyBURL挖掘的规则依赖于饱和度阈值的设置，而学习合理的阈值往往是很难的，导致方法的鲁棒性较差。为解决该问题，Meilicke等[53]提出了Reinforce-AnyBURL算法。该算法利用强化学习方法，将规则的可解释性、置信

chinaXiv:202305.00174v1

度和长度作为奖励进行规则学习，并提出对象标识方法，在规则中添加约束条件以去除冗余和重复的规则。

为避免规则聚合受到冗余信息的影响，Ott 等[54]进一步改进了 AnyBURL 方法，提出了一种可迁移的快速非冗余规则挖掘方法 SAFRAN(Scalable And Fast non-redundant Rule Application)。该方法使用了一种称为非冗余噪声聚合方法，在聚合之前检测并聚类冗余规则，从而减轻其负面影响并提高预测性能。SAFRAN 采用最大聚合方法预测实体的置信度得分，然后基于噪声或方法对不同聚类的预测进行进一步聚合。此外，SAFRAN 采用了网格搜索和随机搜索两种方法，来确定基于类型组合的三元组规则最优聚类阈值。相比 AnyBURL，SAFRAN 挖掘规则的质量更高，推理结果也更准确。

表 1 给出了基于归纳逻辑编程的推理方法的相关信息。基于 ILP 的规则学习和知识图谱推理方法是一种“硬”推理方法，挖掘得出的规则通过置信度评分并与阈值进行比较来判定是否正确。但此类方法受阈值设置的影响具有一定的主观性，并且只适用于精确推理，而无法表示不确定性信息。

表 2 基于归纳逻辑编程的推理方法的相关信息

分类	优点	经典方法	年份	代码地址	不足
基于一阶归纳学习	无监督方法，利用有效类之间的关系学习规则，使用 Markov 网络计算置信度	FOIL	1990	/	需要负实例，不适用于开放世界假设，计算复杂度随规则数指数上升
		N-FOIL	2007	/	
		T-FOIL	2007	/	
		KFOIL	2010	/	
		QuickFOIL	2015	/	
基于关联规则挖掘	利用特定算子挖掘规则，根据置信度修剪规则，约束规则长度、类型以减小搜索空间	AMIE	2013	https://github.com/lajus/amie	需要遍历关系路径，不适用于大规模知识图谱
		AMIE+	2015	https://github.com/lajus/amie	
		AMIE3	2020	https://github.com/lajus/amie	
		RDF2Rules	2015	/	
基于关系路径采样	对目标关系，只对知识图谱中与该关系相关联的三元组进行采样，然后利用采样信息进行规则学习	RuleN	2018	/	只能执行精确推理，无法应用于模糊推理
		AnyBURL	2019	http://web.informatik.uni-mannheim.de/AnyBURL	
		Reinforce AnyBURL	2020	https://web.informatik.uni-mannheim.de/AnyBURL	
		SARFRAN	2021	https://github.com/OpenBioLink/SAFRAN	
		C-NN	2019	/	

4 基于概率图模型和规则的知识图谱推理

chinaXiv:202305.00174v1

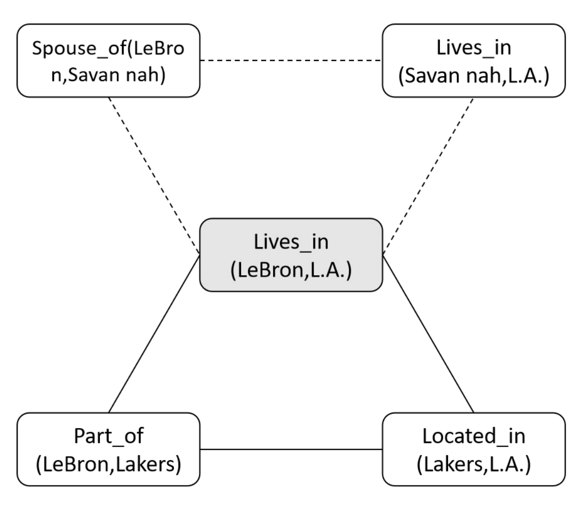
为实现模糊推理，研究人员将概率图模型与规则学习相结合，提出了多种适用于模糊推理的“软”推理方法，主要包括基于马尔可夫逻辑网络的方法、基于推理树的方法和基于概率软逻辑的方法。

4.1 基于马尔可夫逻辑网络的方法

Richardson 等[56]提出的马尔可夫逻辑网络(Markov Logic Network, MLN)是根据预先定义的规则和以知识图谱为单位的事实建立一个概率图形模型，然后学习不同规则的权重。给定任何规则集合$\{\gamma_i\}$，集合中的规则都在知识图谱中对应实例。根据$\{\gamma_i\}$，可以建立一个马尔可夫逻辑网络。

MLN 建模的主要步骤是：(1)为每条规则中的基本原子构建一个节点，如果原子在知识图谱中被观察，则节点的值设为 1，否则设为 0；(2)两个基本原子可以同时用于实例化至少一个规则时，在两个节点之间建立边；(3)所有基本原子形成一个团，对应一个特征，如果基本规则为真，值为 1，否则为 0。

其中，$n_i(x)$是规则γ_i的真实基本原子数，w_i是规则γ_i对应的权值。然后，在 MLN 中应用马尔可夫链蒙特卡洛(Markov Chain Monte Carlo, MCMC)算法进行推理，并通过优化伪似然测度有效地学习权重w_i。根据w_i，给定相邻的基本原子，可以推断出 Lives_in(LeBron,L.A)的概率。



γ_1:Spouse_of(X,Y)∧Live_in(Y,Z)→ Lives_in(X,Z)
γ_2:Part_of(X,Y)∧Located_in(Y,Z)→ Lives_in(X,Z)

图 5 两条规则对应的 MLN 示例

MLN 是第一个被开发用于知识图谱模糊推理的模型。然而，由于三元组之间的图结构复杂，MLN 的推理过程困难且效率低下。此外，知识图谱中缺失的三元组也会通过规则影响推理结果。

4.2 基于推理树的方法

选择性线性确定(Selective Linear Definite, SLD)[57]是一种为查询构造推理树来进行知识图谱推理的手段。相比 MLN，SLD 推理树能够有效缓解三元组结构复杂的问题，具有更高的推理效率。标准 SLD 方法以自顶向下的方式构建推理树，如图 6 所示。它首先利用查询来初始化根节点，然后通过应用每个子句及其实例来递归地创建子目标。

基于 SLD 推理树，Raedt 等[57]提出了 ProbLog 方法，它是逻辑编程方法 Prolog 的概率扩展。与 Prolog 相比，ProbLog 为每个子句c_i增加了一个概率p_i，c_i表示

chinaXiv:202305.00174v1

一个规则或一个基本原子，用于派生查询的示例，最后通过基于推理树的递归来获得推理结果。

ProbLog 中的所有规则或原子都被概率化，例如：0.7:Spouse_of(X,Y)∧Lives_in(Y,Z)→Lives_in(X,Y)；0.8:Part_of(X,Y) ∧Located_in

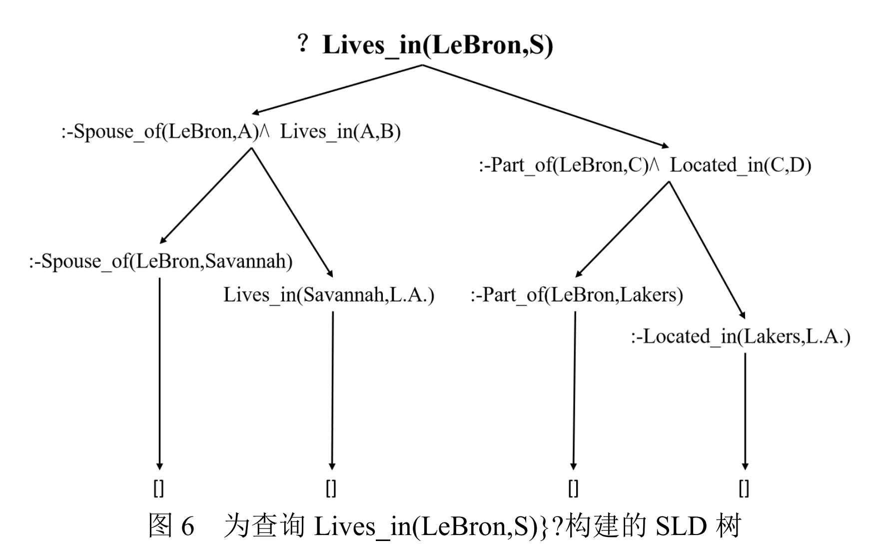


图 6 为查询 Lives_in(LeBron,S)}?构建的 SLD 树

(Y,Z)→Lives_in(X,Y)；1.0:Spouse_of(LeBron,Savan nah)。给定任何一个查询 q，成功概率 $P(q|T)$ 被分解为计算所有可能子句集合的联合概率总和，即：

$$P(q|T)=\sum_{L\subseteq L_T}P(q,L|T)=P(q|L)P(L|T)=P(q|L)\left[\prod_{c_i\in L}p_i\prod_{c_i\in L_T\setminus L}(1-p_i)\right], \quad (6)$$

其中 $T=\{p_1:c_1,\dots,p_n:c_n\}$ 描述了原因 L 的概率分布，$L\subseteq L_T=\{c_1,\dots,c_n\}$。$P(q|L)$ 表示在给定原因集 L 的情况下查询 q 的概率，如果至少有一个答案能实例化 L 并使查询为真，则其值为 1，否则为 0。为了计算 $P(q|L)$，ProbLog 利用 SLD 为 q 构造一个推理树，首先根据查询初始化根节点，然后应用每个子句及其实例来递归地创建子目标，达到结束条件(找到答案或达到最大的树深度)时停止。

为提升推理效率，Cussens[60]提出了随机逻辑程序(Stochastic Logic Programs, SLP)方法，SLP 是带有参数化子句的逻辑程序，它定义了目标反演的对数线性分布，而对数线性分布则通过边缘化提供了绑定于变量的分布，使得 SLP 能够表示复杂的规则分布。SLP 定义了一个遍历 SLD 树的随机过程，定义在节点上的概率分布通过向上加权所需的答案子句和向下加权其他子句来学习。此外，Cussens 还提出了一种故障调整最大化(Fault-Adjusted Maximisation, FAM)方法，FAM 是 EM 算法的一个实例，专门应用于标准化 SLP，并在迭代最大化方法中为计算参数更新提供了一种封闭形式，能够有效地提升规则优化的速度。

Wang 等[61]利用 PPR(Personalized PageRank)算法的偏差抽样策略替代 SLP 的随机抽样策略，提出了 ProPPR 方法。与式 6 中直接为 $P(L|T)$ 设置一个概率不同，ProPPR 使用 PPR 中基于一些预定义的特征来计算每个子句的概率，然后基于局部分区方法，设计了一个局部基本过程以实现推理。这个局部基本过程具有的查询能力能将规则权重学习任务划分为多个单独的梯度计算，实现并行学习，以提升规则学习和推理的效率。此外，实验表明，即使没有权重学习，ProPPR 在实体解析任务上也表现良好，而监督权重学习提高了准确性。

chinaXiv:202305.00174v1

除了 ProbLog, SLP 和 PPR 之外，还有 Datalog[62]、MV-Datalog±View[63]等方法也采用了类似于 SLD 的推理逻辑。基于推理树的方法尽管相比 MLN 在速度方面有所改进，但面对大量缺失三元组的知识图谱时，推理精度和速度都不理想。

4.3 基于概率软逻辑的方法

针对三元组缺失的知识图谱推理问题，Kimmig 等[64]提出了概率软逻辑(Probabilistic Soft Logic, PSL)。PSL 是一种在关系领域中进行集体概率推理的框架，由一组一阶逻辑规则、连接体(body)和单个字面头(head)组成。

类似于 ProbLog 方法，PSL 的规则被赋予了权重。为了确定满足基本规则的程度，PSL 分别使用 Lukasiewicz t-范数[65]及其相应的协范数作为逻辑 AND 和 OR 的松弛：

$$\begin{aligned}&\ell_1 \tilde{\wedge} \ell_2 = \max\left\{0, I(\ell_1)+I(\ell_2)-1\right\}\\&\ell_1 \tilde{\vee} \ell_2 = \min\left\{I(\ell_1)+I(\ell_2),1\right\}\\&\tilde{\neg}\ell_2 = 1-I(\ell_1)\end{aligned} \quad, \tag{7}$$

“~”表示自布尔域的松弛，I 表示给定的解释，ℓ 表示原子，通过上式，布尔真值松弛为区间 [0， 1] 中的连续软真值。对于一个基本规则 $r \equiv r_{body} \rightarrow r_{head} \equiv \tilde{\neg} r_{body} \tilde{\vee} r_{head}$，对 r 中原子的解释 I 决定了 r 是否满足要求，则利用 $d_r = \max\left\{0, I(r_{body}) - I(r_{head})\right\}$ 来计算规则距离满足要求的“差距”。通过松弛和距离计算，PSL 将相似性函数集成到图模型中。此外，PSL 还将一阶公式的语法限制为具有连接体规则的语法。这两个特性使得推理转化为连续空间中的凸优化问题，可以在知识缺失的情况下进行规则学习，求解更加方便。

Pujara 等[66]为减少知识图谱识别和推理时间，实现在信息系统噪声背景下的知识图谱推理，提出了一种本体感知分割方法 OAP(Ontology-Aware Partitioning)。OAP 建立以实体和关系为节点、本体约束关系为边的本体知识图，如图 7 所示，并使用 PSL 来定义知识图上的联合概率分布。然后，利用边缘最小分割技术，来划分本体中存在的关系和标签，进而创建关系和标记特定实例提取的相应分区，执行知识推理。针对本体论信息的贡献和与数据中不同关系、标签不一致导致的分区不平衡问题，OAP 根据关系或标签的提取频率对每个顶点进行加权，约束图切割以产生具有相等顶点权重的簇。OAP 利用本体进行分区提取，保留联合方法构建知识图的优势，并显著提高运行效率，能够从信息提取系统的噪声输出中联合推理知识图谱。

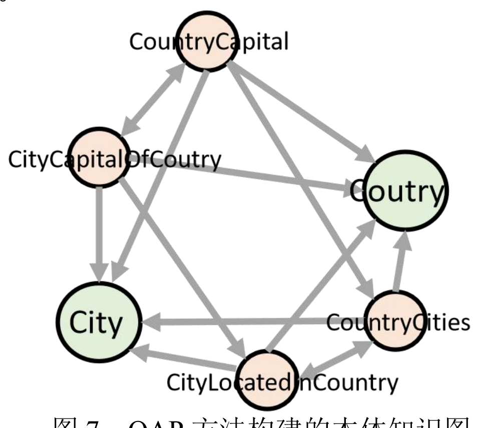


图 7　OAP 方法构建的本体知识图

chinaXiv:202305.00174v1

为适应大规模复杂知识图谱，Bach 等[67]提出了铰链损失马尔可夫随机场(Hinge-Loss Markov Random Field, HL-MRF)模型。HL-MRF 是一种结合随机算法、概率图形模型和模糊逻辑社区的概率图模型。首先，研究者们证明了 MAX SAT 的随机算法、使用布尔逻辑定义的离散 MRF 的局部一致性松弛和模糊逻辑连续信息推理三种方法可以统一到相同的凸优化目标上，然后将推理目标推广为 Hinge-Loss 特征的加权和。HL-MRF 用铰链损失势和逻辑规则之间的联系来定义 PSL，使得模型很容易扩展到大型知识图谱中。此外，PSL 提供的语法使模型能够更好地被用户理解和拓展。

表 2 给出了基于概率图模型和规则的推理方法的相关信息。尽管能够实现模糊推理，但由于此类方法本质上还是依赖于规则的搜索挖掘，需要在知识图谱上进行复杂的遍历计算，因此在大规模图谱中的精度、效率和鲁棒性都不强，目前未能在大规模知识图谱中得到广泛的应用。

5 基于嵌入表示和规则的知识图谱推理

传统的基于显示规则学习的知识图谱推理方法具有良好的可读性和可解释性，但这些有限和离散的规则往往难以完整描述数据之间的所有内在关系，对模糊和噪声也不具有鲁棒性。知识图谱嵌入表示将实体和关系表示为低维连续向量空间中的稠密向量，可以有效表示和度量实体、关系间的语义关联，提升计算效率，缓解数据稀疏，对模糊和噪声更加鲁棒。如何有效结合嵌入表示和规则学习的优点提升推理能力，已经成为学者们的重点研究内容之一。在现有的研究工作，

表 2 基于概率图模型和规则的推理方法的相关信息

分类	特点	经典方法	年份	代码地址	不足
基于马尔可夫逻辑网络	根据规则和以知识图谱为单位的事实建立一个概率图形模型，然后学习不同规则权重	MLN	2005	https://github.com/danielnyga/pracmln	推理过程困难且效率低下，推理准确度受缺失三元组影响
基于推理树	基于选择性线性确定方法 SLD，以自顶向下的方式构建推理树，通过查询和子句、实例递归创建子目标	ProbLog	2007	/	面对大量缺失三元组的知识图谱时，推理精度和速度不佳
		SLP	2001	/	
		ProPPR	2013	/	
基于概率软逻辑	通过松弛和距离计算将相似性函数集成到图模型中，把推理转化为连续空间凸优化问题，可在知识缺失的情况下学习规则。	PSL	2012	http://psl.umiacs.umd.edu	在大型图谱中效率不佳，鲁棒性差。
		OAP	2013	https://github.com/linqs/KnowledgeGraphIdentification	
		HL-MRF	2017	http://psl.linqs.org	

chinaXiv:202305.00174v1

chinaXiv:202305.00174v1

基于嵌入表示和规则的知识图谱推理方法可以分为规则嵌入表示学习的推理方法和嵌入表示与规则相互增强的推理方法。

5.1 规则嵌入表示学习的方法

此类方法通过规则嵌入或规则增强嵌入表示进行推理。规则可以为知识图谱嵌入表示提供信息，并通过规则推理出的三元组为嵌入表示提供更多训练集，提升嵌入表示的效率和质量，同时也可以为推理结果提供更多的解释依据，增强推理可解释性。

Guo 等[68]于 2016 年提出了 KALE 方法，这也是结合嵌入表示和规则进行知识图谱推理的最早研究之一。KALE 将规则和三元组嵌入到一个统一框架中，对三元组和规则进行联合嵌入建模。方法将三元组视作基本原子，给定一个逻辑规则，首先用集中的实体进行实例化，将规则视为由基本原子和逻辑连接符构成的复杂公式。对三元组的建模采用 TransE[69]方法，将其真值定义为：

$$I\left(e_i,r_k,e_j\right)=1-\left\|\frac{1}{3\sqrt{d}}\mathbf{e}_i+\mathbf{r}_k-\mathbf{e}_j\right\|, \tag{8}$$

d 为嵌入维度。而对于规则建模，则采用模糊逻辑的方法，考虑两种规则：一种是 $\forall x,y:(x,r_s,y)\rightarrow(x,r_t,y)$，记作 $f_1\triangleq(e_m,r_s,e_n)\rightarrow(e_m,r_t,e_n)$；另一种是 $\forall x,y,z:(x,r_{s1},y)\wedge(y,r_{s2},z)\rightarrow(x,r_t,z)$，记作 $f_2\triangleq(e_l,r_{s1},e_m)\wedge(e_m,r_{s2},e_n)\rightarrow(e_l,r_t,e_n)$，两者的真值 I 的计算方法为：

$$\begin{aligned}&I\left(f_1\right)=I\left(e_m,r_s,e_n\right)\cdot I\left(e_m,r_t,e_n\right)-I\left(e_m,r_s,e_n\right)+1,\\&I\left(f_2\right)=I\left(e_l,r_{s1},e_m\right)\cdot I\left(e_m,r_{s2},e_n\right)\cdot I\left(e_l,r_t,e_n\right)-I\left(e_l,r_{s1},e_m\right)\cdot I\left(e_m,r_{s2},e_n\right)+1,\end{aligned} \tag{9}$$

e 表示实体，r 表示关系，I 的值越大，规则被满足的概率越大。最后，最小化可训练公式集合上的一个全局损失，以学习与三元组和规则都相容的嵌入表示，即：

$$\begin{aligned}&\min_{\{\mathbf{e}\}\{\mathbf{r}\}}\sum_{f^+\in\mathcal{F}}\sum_{f^-\in\mathcal{N}_{f^+}}\left[\gamma-I\left(f^+\right)+I\left(f^-\right)\right]_+,\\&\text{s.t.}\|\mathbf{e}\|_2\le1,\forall e\in\mathcal{E};\|\mathbf{r}\|_2\le1,\forall r\in\mathcal{R},\end{aligned} \tag{10}$$

$\mathbf{e},\mathbf{r}$ 表示实体和关系的嵌入向量，$\mathcal{E},\mathcal{R}$ 表示实体和关系空间，γ 表示允许的边界距离。

Wang 等[70]借鉴 ProPPR 方法，提出了一种基于矩阵分解进行一阶逻辑嵌入的方法 ProPPR-MF(ProPPR Matrix Factorization)。该方法首先提出基于结构梯度的结构学习策略，从事实中生成合理的推理公式；然后，利用背景事实、训练示例和推理公式来构建有根据的证明图；最后，将训练示例映射到二元矩阵的行中，将推理公式映射到二元矩阵的列中，使用可迁移矩阵分解方法，通过低秩近似学习实例和逻辑公式的潜在连续表示。ProPPR-MF 的步骤如图 8 所示。通过逻辑规则嵌入学习，将离散的逻辑事实和谓词表示为连续的矩阵向量，提升了推理的效率，使得方法能够扩展应用到大型复杂知识图谱中。

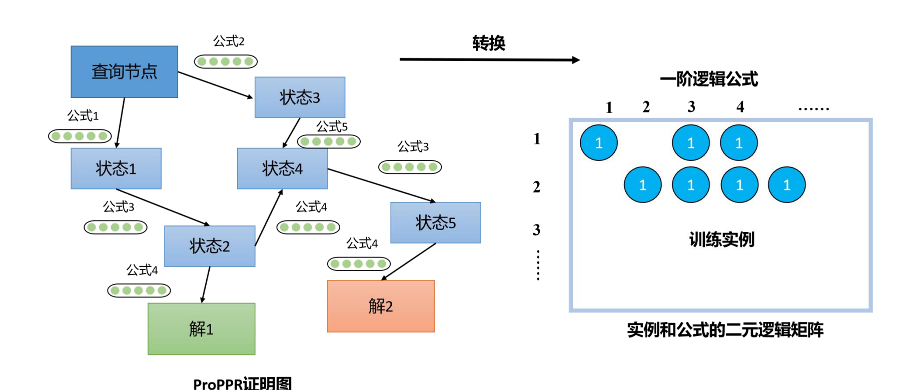


图 8　ProPPR-MF 的方法框架

传统的规则嵌入学习方法采用一次性注入逻辑规则，忽略了嵌入学习与逻辑推理之间的交互特性。这些方法只关注硬规则，而硬规则通常需要大量的手工工作来创建或验证。为克服该问题，在 KALE 的基础上，Guo 等人进一步提出了基于软规则引导的方法 RUGE[71]，利用迭代方式替代一轮规则注入。RUGE 利用 AMIE 获取规则，但不是直接将基本规则视为正实例，而是将规则导出的三元组视作未标记三元组，利用可观测三元组更新实体和关系嵌入。由于未标记三元组的真实性难以确定，因此基于当前嵌入来预测每个未标记三元组的概率。最后，基于标记的和未标记的三元组来更新嵌入。三元组的嵌入采用 ComplEx 方法[72]，规则嵌入则与 KALE 采用相同的 t-norm 模糊逻辑方法，其迭代学习过程的算法伪代码如算法 1 所示。

KALE 和 RUGE 依据规则或公式组成部分的分数组合的计算其评分，但由于三元组的分数是分别估计的，因此即使规则或公式中的三元组彼此完全无关，也可能导致该规则或公式的分数较高。为了解决这一问题，Wang 等[73]将一个三元组或一个基本规则转换为一阶逻辑，然后根据一阶逻辑中包含的实体和关系的

算法 1　RUGE 的迭代学习过程

Require: 标记三元组 $\mathcal{L}=\{(x_l,y_l)\}$,未标记三元组 $\mathcal{U}=\{(x_u)\}$,一阶逻辑规则 $\mathcal{F}=\{(f_p,\lambda_p)\}$ f_p为规则，λ_p为置信度 以及其基本原子 $\mathcal{G}=\{(g_{pq})\}$

1.随机初始化实体和关系嵌入 $\Theta^{(0)}$

2.**for** $n=1:N$ **do**:

3.　从 $\mathcal{L}/\mathcal{U}/\mathcal{G}$ 中抽取 minibatch $\mathcal{L}_b/\mathcal{U}_b/\mathcal{G}_b$

4.　$S_b \leftarrow \text{SoftLabelPrediction}\left(\mathcal{U}_b,\mathcal{G}_b,\Theta^{(n-1)}\right)$　#标签预测

5.　$\Theta^{(n)} \leftarrow \text{EmbeddingRectification}\left(\mathcal{L}_b,\mathcal{U}_b,\mathcal{G}_b\right)$　#嵌入更新

6.**end for**

Ensure: $\Theta^{(N)}$

嵌入进行向量或矩阵运算，以对这个一阶逻辑进行评分。表 3 说明了一阶逻辑的格式，表 4 介绍了如何通过数学表达式对一阶逻辑进行评分。通过上述转化，包含在同一规则中的不同三元组在向量空间中直接交互，确保规则及其编码格式都具有一对一的映射转换，对规则的打分更加合理，推理的准确度也更高。

多数基于嵌入式表示和规则的方法忽略逻辑规则的传递性和反对称性，针对该问题，Wang 等[74]提出了 TARE 方法，其核心思想是将知识图谱三元组、已有关系和逻辑规则相结合，以学习知识图谱嵌入，利用逻辑规则的传递性和反对称性，对逻辑规则中的关系类型进行近似排序。首先，采用 ComplEx 方法嵌入实体和

chinaXiv:202305.00174v1

关系，关系向量嵌入被限制为非负的，并将问题简化为非负矩阵分解，嵌入学习过程损失函数为：

$$\mathcal{L}_K = \min_\Theta \sum \log\left(1+\exp\left(-Y_{sro}\left(\phi\left(s,r,o\right)\right)\right)\right)+\lambda\Theta_2, \quad (11)$$

其中$\phi(s,r,o)$是 ComplEx 方法中的能量函数，Θ为模型参数，$s,o\in\mathcal{E},r\in\mathcal{R}$，$\{Y_{sro}\}_{(s,r,o)\in\Omega}\{-1,1\}$，$\Omega\in\mathcal{E}\times\mathcal{R}\times\mathcal{E}$，$|\Omega|$是三元组为真或假的标签集合。然后分别对复数向量嵌入的实部和虚部进行排序获得关系之间的顺序。例如，对于规则$r_a\wedge r_b\Rightarrow r_c$，排序过程的损失函数为：

$$\mathcal{L}_R = \min \sum_{(r_a,r_b,r_c\in P)} F\left(r_a,r_b,r_c\right)+\sum_{(r'_a,r'_b,r'_c\in N)} \max\left(0,\alpha-F\left(r'_a,r'_b,r'_c\right)\right), \quad (12)$$

表 3 三元组及其基本规则的对应格式

三元组及基本规则	格式
(h,r,t)	$r(h)\Rightarrow t$
$(h,r_1,t)\Rightarrow(h,r_2,t)$	$\left[(h\in C)\wedge\left[r_1(h)\Rightarrow t\right]\right]\Rightarrow\left[r_2(h)\Rightarrow t\right]$
$(e_1,r_1,e_2)+(e_2,r_2,e_3)\Rightarrow(e_1,r_3,e_3)$	$\left[r_1(e_1)\Rightarrow e_2\right]\wedge\left[r_2(e_2)\Rightarrow e\right]\Rightarrow\left[r_3(e_1)\Rightarrow e_3\right]$
$(h,r_1,t)\Leftrightarrow(t,r_2,h)$	$\left[\left[r_1(h)\Rightarrow t\right]\Rightarrow\left[r_2(t)\Rightarrow h\right]\right]\wedge\left[r_2(t)\Rightarrow h\right]\Rightarrow\left[r_1(h)\Rightarrow t\right]$

表 4 一阶逻辑及其数学表达式

一阶逻辑	数学表达式
$r(h)$	$\mathbf{r}+\mathbf{h}$
$a\Rightarrow b$	$\mathbf{a}-\mathbf{b}$
$h\in C$	$\mathbf{h}\bullet\mathbf{C}$（$\mathbf{C}$ 为矩阵）
$a\wedge b$	$\mathbf{a}\bullet\mathbf{b}$
$a\Leftrightarrow b$	$(\mathbf{a}-\mathbf{b})\bullet(\mathbf{a}-\mathbf{b})$

P和N表示逻辑规则的正反集合，α为边缘超参数，F表示逻辑规则的惩罚函数得分，r'_a,r'_b,r'_c为将结果中的r_c替换为随机关系$r\in R$来构造的反规则集合。最后，对损失函数$\mathcal{L}=\mathcal{L}_K+\mathcal{L}_R$，利用随机梯度下降法进行联合优化。TARE 方法中，逻辑规则直接合并到关系类型表示中，而不是用具体的实体实例化。因此，学习到的嵌入不仅与三元组兼容，而且与规则兼容，并且关系类型的嵌入是近似有序的。

Ding 等[75]提出了 ComplEx-NNE-AER 方法，该方法将实体表示的非负约束(No-NEgative Constraints, NNE) 和关系表示的近似包含约束 (Approximate Entailment constraints on Relation, AER)融入嵌入模型，以在不影响嵌入和可迁移性的前提下，在嵌入空间的结构中增加先验信念。NNE 有助于学习实体的紧凑和可解释表示，AER 进一步将关系之间蕴含的逻辑规则编码到它们的分布式表示中。该方在 ComplEx 方法的基础上施加 3 条约束进行嵌入表示学习，即：

chinaXiv:202305.00174v1

$$
\begin{aligned}
&0 \le \mathbf{Re}(\mathbf{e}), \mathbf{Im}(\mathbf{e}) \le 1, \forall e \in \mathcal{E} \\
&\lambda\left(\mathbf{Re}(\mathbf{r}_p) - \mathbf{Re}(\mathbf{r}_q)\right) \le \boldsymbol{\alpha}, \\
&\lambda\left(\mathbf{Im}(\mathbf{r}_p) - \mathbf{Im}(\mathbf{r}_q)\right)^2 \le \boldsymbol{\beta},
\end{aligned} \tag{13}
$$

其中 **Re** 和 **Im** 分别表示嵌入向量的实部和虚部，$r_p \xrightarrow{\lambda} r_q$ 表示关系 r_p 以 λ 的置信度蕴含 r_q，$\boldsymbol{\alpha}$ 和 $\boldsymbol{\beta}$ 为松弛变量。式 13 确保实体表示中只存储正性属性，并将实体表示保持在 $[0,1]^d$ 的超立方体内；式 13 确保了模型无需遍历所有潜在实体对，并利用不同置信度水平的编码使得模型对不确定性更加鲁棒。ComplEx-NNE-AER 方法简单有效，普适性强，NNE 和 AER 提高了模型的效率和可解释性，增加了嵌入空间的结构，使其能够适用于大规模知识图谱。

Niu 等[76]将规则学习、嵌入表示和路径学习相结合，提出了一种基于规则和路径的联合嵌入方法 RPJE(Rule and Path-based Joint Embedding)。首先，从知识图谱中挖掘不同长度(规则体包含关系的数量)的 Horn 规则，并编码用于嵌入表示。然后，利用长度为 2 的规则对路径进行精确组合，使用长度为 1 的规则创建关系之间的语义关联，约束关系嵌入。此外，优化时还考虑了每个规则的置信度，以保证将规则应用于嵌入表示的可用性。RPJE 的模型结构如图 9 所示。该方法充分利用了逻辑规则的可解释性和准确性、知识图谱嵌入的泛化性以及路径的补充语义结构，可以有效缓解嵌入表示学习方法难以适应稀疏知识图谱和可解释性差的问题。

此外，Tang 等提出了 RulE[77]方法，通过嵌入表示建模逻辑规则和三元组。RulE 将实体、关系和逻辑规则统一嵌入空间中进行共同表示，然后学习每个逻辑规则的嵌入。该方法以软方式执行逻辑规则推理，并计算每个基本规则的置信度分数。相比单纯的嵌入表示方法， RulE 允许将先验逻辑规则信息注入嵌入空间，这提高了知识图嵌入的泛化能力。此外，学习到的规则的置信度分数通过软控制每个规则的贡献来改善逻辑规则推理过程，从而缓解了规则的鲁棒性问题。

规则嵌入表示学习的方法能够提升知识推理的效率和可解释性，但由于此类

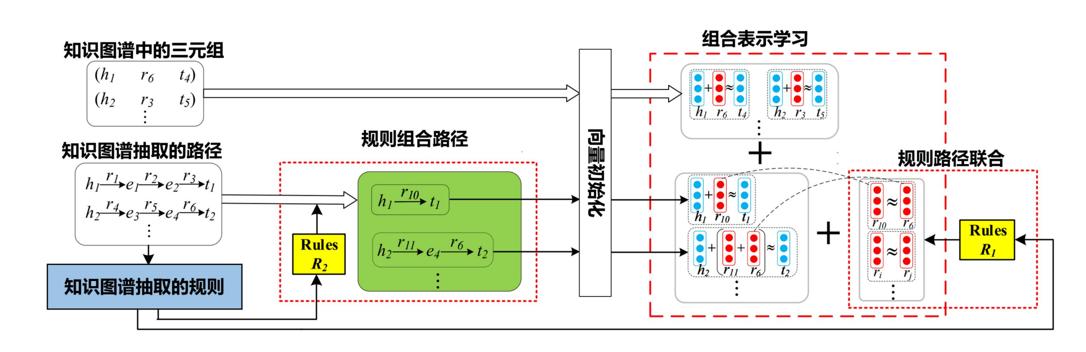


图 9　RPJE 的模型结构

方法本质上还是对知识图谱进行嵌入表示，其模型整体本身依然是不可解释的。换句话说，此类方法只具有结果可解释性，而不具有过程可解释性。

5.2　嵌入表示与规则相互增强的知识图谱推理

规则学习和嵌入学习是相辅相成的，一方面，规则可以为稀疏知识图谱推断额外的三元组，实现更准确的嵌入学习；另一方面，包含丰富语义编码的嵌入表

chinaXiv:202305.00174v1

示可以将离散图搜索的规则学习转化为向量空间的计算，从而显著缩小搜索空间。因此，嵌入表示与规则相互增强的方法可以兼顾效率提升和推理过程的可解释性。

Zhang 等[78]认为，现有的规则学习增强知识图谱嵌入的方法只利用了已有的规则，未能在推理过程中利用嵌入表示对规则进行更新，这会对推理精度造成较大影响。针对此问题，他们提出了 IterE 方法。IterE 以同时学习嵌入和规则为目标，在每次迭代中根据更新的嵌入推断出新的规则。IterE 算法主要包括三个部分：(1)嵌入学习；(2)公理归纳；(3)公理注入。嵌入学习步骤主要是利用知识图谱中存在的三元组和规则推断的三元组学习实体和关系的嵌入；公理归纳步骤利用一种有效修剪策略生成公理池，然后根据线性映射假设的规则结论，对关系嵌入进行计算，为池中的每个公理打分；公理注入部分主要是利用公理的演绎能力，根据基本公理推断出关于稀疏实体的新三元组，并将这些新三元组注入到知识图谱中，以改进稀疏实体嵌入。IterE 的模型框架如图 10 所示。IterE 利用公理提高了稀疏嵌入的质量，利用嵌入提高了规则学习的效率和质量，而迭代训练则确保了更好的链路预测性能和高质量的规则学习结果。

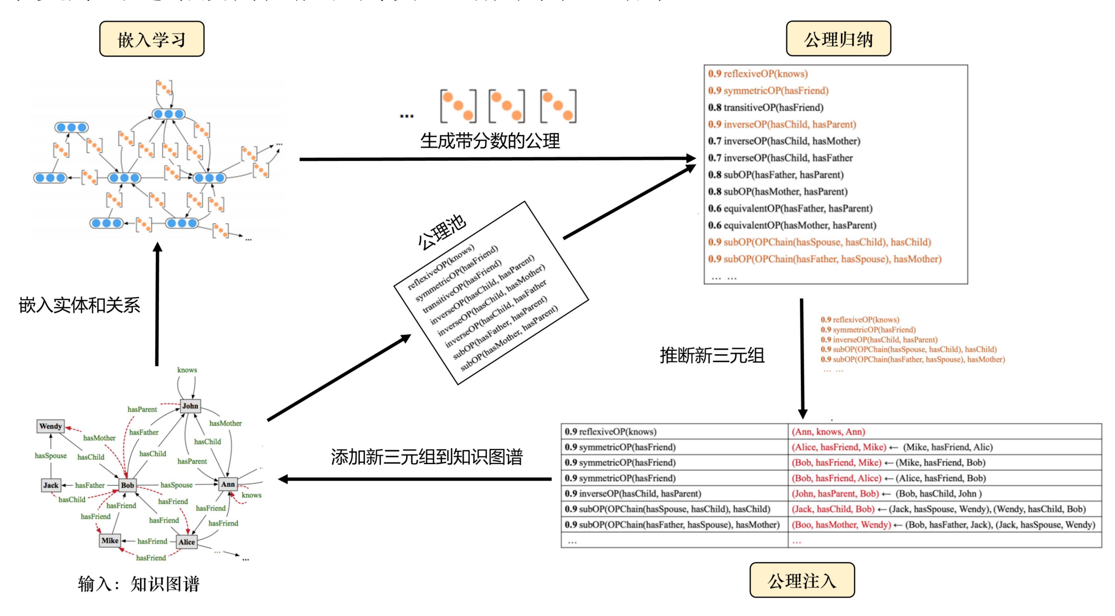


图 10　IterE 的模型框架

由于知识图谱中缺失事实的存在，仅仅利用其本身的信息学习到高质量的规则是困难的。针对这一问题，Ho 等[79]提出了一种在三元组缺失的背景下借助嵌入模型反馈来学习高质量规则的方法 RuLES。该方法从一个初始的规则开始，逐个在规则体中增加原子，并利用联合知识图谱和文本语料库训练的嵌入模型为规则生成的三元组进行质量评分，以修剪和更新规则。RuLES 方法的步骤如图 11 所示。

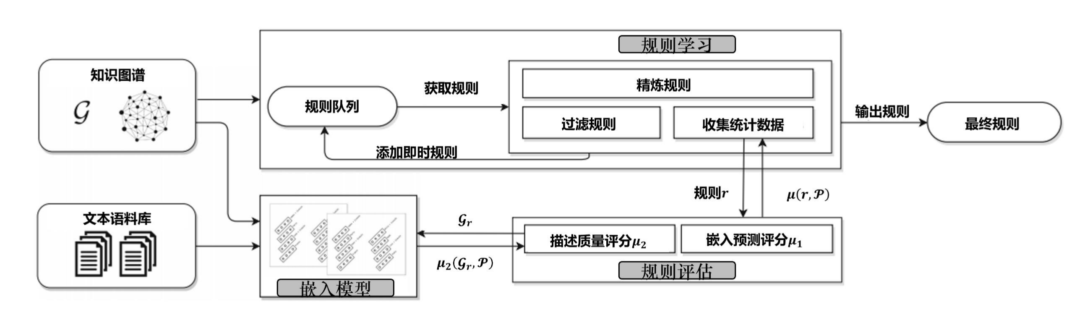


图 11　RuLES 方法步骤

chinaXiv:202305.00174v1

RuLES 的模型输入为知识图谱$\mathcal{G}$、文本语料库和用于终止规则构造的指定参数，包括规则学习、嵌入学习和规则评估三个模块。规则学习模块在以知识图谱为依据构建规则，规则评估模块使用$\mathcal{G}$和嵌入模型给出的质量评分函数f为规则r提供质量分数μ(包括基于规则的描述质量评分μ_1和基于嵌入的预测概率评分μ_2)，嵌入模块利用$\mathcal{G}$和文本计算，$\mathcal{P}=(\mathcal{G},f)$称为“概率”知识图谱。模型将规则建模为原子序列，其中第一个原子是规则头，其他原子是规则体。初始阶段，出现在$\mathcal{G}$中的所有可能的二元原子都被添加到具有空规则体的规则队列中。在每次迭代中，从队列中选择一条规则。若规则满足过滤条件，则系统将其作为输出返回；若规则未被过滤，则使用精炼操作进行规则扩展，将新候选规则加入队列。通过在不完整图谱中综合评估规则质量和嵌入反馈的规则质量，RuLES 能学到更加高质量的规则以用于推理。

为提升谓词数量较多的大规模知识图谱推理精度，Omran 等[80]提出了 RLvLR 方法。该方法结合嵌入表示学习技术和一种新的抽样方法，从大型知识图中学习规则。RLvLR 主要应用了三种技术：(1)子图采样技术，通过删除与目标谓词相关度较小的三元组来缩小输入知识图谱的规模，提升嵌入表示的可迁移性；(2)参数嵌入技术，对规则进行质量评估；(3)基于矩阵运算的规则质量评估技术，以避免昂贵的查询应答操作。首先，RLvLR 将目标谓词周围路径长度小于N的实体及关系组织成子图；然后对子图进行搜索，查找与谓词相关的路径以发现规则，采用 RESCAL 方法[125]为规则中的实体和谓词生成嵌入表示，并基于语义相似和谓词共现的思想设计评分函数，指导规则修剪，得到候选规则集合；最后，通过有效的邻接矩阵乘法进行最终评估，得到一套封闭规则。在 RLvLR 的基础上，Orman 等还提出了一种从知识图谱流中学习规则新方法 RLvLR-stream，用于完成时序知识图谱的预测。RLvLR 在规则学习的质量和效率方面性能突出，链路预测的精度也优于同类方法。算法 2 给出了 RLvLR 算法的具体流程。

在大规模知识图谱中，基本规则的数量较多，IterE 和 RuLES 等嵌入表示与规则相互增强的方法需要对基本规则进行采样来解决可迁移性问题，这对推理效率产生了很大影响。针对这个问题，Cheng 等[81]提出了推理框架 UniKER，将逻辑规则限制为 Horn 规则，充分利用规则中的知识，使规则学习和嵌入表示以一种高效的方式相互增强。为确定 Horn 规则，UniKER 设计了一种改进的前向链

算法 2　RLvLR 的规则学习

Require：知识图谱$\mathcal{KG}$，目标谓词P，最大规则长度l

1. $\mathcal{KG}'=\text{Sampling}(\mathcal{KG},P)$
2. $(\mathcal{P},\mathcal{A})=\text{Embedding}(\mathcal{KG}')$
3. $R':=\varnothing$

4.**for** $2\le k\le l$ **do**:

5. 　添加 $\text{PathFinding}(\mathcal{KG}',P,\mathcal{P},\mathcal{A},k)$ 至 R' 中

6.**end for**

7. $R'=\text{Evaluation}(R',\mathcal{KG})$

Ensure:一套封闭路径规则R

迭代算法，该算法采用了“惰性推理”，只需要用到基础谓词和规则中的一小部分，并能在推理进行时激活更多的基础谓词和规则，避免了大量从未使用过的基础谓词和规则的计算，极大地提高了推理效率。然后，通过基于嵌入的变分推理对知识图谱进行更新，利用嵌入表示方法添加有用的隐藏三元组，删除知识图谱和推

chinaXiv:202305.00174v1

断结果中的不正确三元组。UniKER 充分利用了规则中包含的知识以实现更好的嵌入，同时嵌入表示通过提供更多潜在可用的三元组来增强正向链接，使得规则学习和嵌入表示得以相互促进，共同提升推理质量。

嵌入表示与规则相互增强的方法以链路预测和规则学习为目标，可以更快地学习到更高质量的规则，有效增强推理结果的可解释性。但随着推理跳数的增加，预测性能对知识图谱的稀疏性更加敏感，推理效率也有所下降。表 5 给出了基于嵌入表示和规则的推理方法的相关信息。

6 基于神经网络和规则的知识图谱推理

近年来，深度学习技术在计算机视觉、自然语言处理和网络链接预测等领域任务中都取得了优异的性能，也已经被广泛用于知识图谱的表示学习和推理中。深度学习方法通过建立各种神经网络模型，能够对数据内部的隐式特征和相关性进行有效建模，在速度、精度、鲁棒性和泛化性方面表现优越，但无法提供对推理结果的解释。融合神经网络和规则以实现高效、可解释的推理，已经成为知识图谱推理的研究重点。根据神经网络和规则在推理中何者占主导作用，可以将此类方法分为规则改进神经网络建模的知识图谱推理和神经网络辅助规则学习的知识图谱推理。

6.1 规则改进神经网络建模的知识图谱推理

规则改进神经网络建模的知识图谱推理主要聚焦如何利用逻辑规则提升神经网络嵌入和预测的能力。在这类方法中，神经网络不仅学习知识图谱中原始观察到的三元组上的实体和关系嵌入，还学习根据一些预定义规则推断得到的三元组或基本规则。规则可以指导神经网络的训练过程，而添加的三元组可以作为训练数据集的增广，这使得基于神经网络的知识图谱推理得到的结果更加具有可解释性，且模型的效率也有所提高。

Wang 等[82]提出了逻辑注意力网络(Logic Attention Network, LAN)，通过聚合基于规则和基于网络的注意力权重的邻居来进行实体嵌入。他们首先提出，一个理想的领域信息聚合器具备三种属性，即置换不变性、冗余感知性和查询关系感知性，并基于该思想设计 LAN。LAN 以基于 RNN[88]的 Encoder-Decoder 为基本框架，使用逻辑规则机制和神经网络机制来联合邻居注意力。逻辑规则机制用于在粗粒度的关系水平上估计权重，将规则之间的潜在依赖关系 $r \Rightarrow q$(q 表示查询关系)量化表示为置信度 $\mathcal{P}(r \Rightarrow q)$，如果 r 是 q 的重要邻居，r 应当满足 $\mathcal{P}(r \Rightarrow q)$ 足够大，$\mathcal{P}(r' \Rightarrow r)$ 较小，其中 r' 表示除了 r 和 q 之外的其他关系。然后，定义逻辑注意力权重：

$$\alpha_{j|i,q}^{\text{Logic}} = \frac{\mathcal{P}(r \Rightarrow q)}{\max\left(\left\{\mathcal{P}(r' \Rightarrow r) \mid r' \in N_{\mathcal{R}}(e_i) \wedge r' \neq r\right\}\right)}, \tag{14}$$

通过计算 $\alpha_{j|i,q}^{\text{Logic}}$，逻辑规则机制解决了查询关系感知和邻域冗余感知强度问题。同时，引入神经网络机制来表示细粒度上的邻居权重，计算描述邻居嵌入后的重要度权重 $\alpha_{j|i,q}^{\text{NN}}$。最后，对两种机制进行加权求和以描述邻居实体重要性：

$$\mathbf{e}_i^O = \sum_{r, e_j \in N_{\mathcal{K}}(i)} \left(\alpha_{j|i,q}^{\text{Logic}} + \alpha_{j|i,q}^{\text{NN}}\right) T_r\left(\mathbf{e}_j^I\right), \tag{15}$$

T_r 表示嵌入函数，$\mathbf{e}_i^O$ 为查询实体嵌入，$\mathbf{e}_j^I$ 为邻居实体嵌入。LAN 对实体的表示学习能力更强，在推理效率、准确度和可解释性方面有明显改进。

chinaXiv:202305.00174v1

Sun 等[83]提出了一种用于在开放领域问答任务中附带文本规则信息的不完整知识图谱推理方法 GRAFT-Nets(Graphs of Relations Among Facts and Text Networks)。该方法主要有两个方面的创新，一是针对知识图谱节点与文本节点，分别提出了异构更新规则，构建 LSTM[84]用于更新文本规则节点之间的传播信息；二是借鉴个性化 PageRank(PPR)[85]算法的思想，提出了一种有向传播方法约束图嵌入的传播方向，使其始终保持在问题的种子节点开始的路径方向。首先，针对查询问题，使用 LSTM 初始化文本规则的潜在表示，然后利用前馈网络分别对实体和文本规则节点进行更新，最后利用关注关系和个性化传播引入与查询问题之间的依赖，前者确保嵌入更多地沿着与问题相关的边传播，后者使得沿着与问题相关的路径获得高权重。GRAFT-Nets 可以有效处理知识图谱和文本的异构图，从包含文本规则信息和实体、关系的特定问题子图中提取答案。

大多数嵌入表示学习方法假定训练时所有测试实体都是可用的，这导致一旦出现新的实体，重新训练嵌入会非常耗时。针对该问题，He 等[86]提出了一种基于图神经网络的嵌入表示框架——虚拟邻居(Virtual Neighbor, VN)网络。VN 网

表 5 基于嵌入表示和规则的推理方法的相关信息

分类	特点	经典方法	年份	代码地址	不足
规则嵌入表示学习方法	规则为稀疏知识图谱推断额外的三元组，提升嵌入表示的效率和质量，也可以为推理结果提供更多的解释依据，增强推理可解释性	KALE	2016	https://aclanthology.org/D16-1019/	本质上是为了学习更好的嵌入表示，其模型整体本身依然是不可解释的
		ProPPR+MF	2016	/	
		RUGE	2018	https://github.com/iieir-km/RUGE	
		Wang et al. [73]	2019	https://github.com/LIBBLE/LIBBLE-MultiThrea	
		TARE	2017	/	
		ComNNE-AER	2018	https://github.com/iieir-km/ComplEx-NNE_AER	
		RPJE	2019	/	
		RulE	2023	https://github.com/XiaojuanTang/RulE	
嵌入表示与规则相互增强的知识图谱推理	利用包含丰富语义编码的嵌入表示可以将离散图搜索的规则学习转化为向量空间的计算，从而显著缩小搜索空间	IterE	2019	https://github.com/wencolani/IterE	随着推理跳数的增加，预测性能对知识图的稀疏性更加敏感，推理效率有所下降
		RuLES	2018	https://github.com/hovinhthinh/RuLES	
		RLvLR	2019	/	
		UniKER	2021	/	

络重点解决三个方面的挑战：首先，为缓解邻居稀疏性，引入规则推断虚拟邻居的概念，即通过约束规则给邻居贴上软标签，而不是简单的将其视为固定不变的；其次，现有的许多方法只使用单跳或双跳邻居进行聚合，忽略了可能有帮助的多跳信息，VN 网络利用图卷积网络，可以同时识别逻辑和对称路径规则来捕获复杂的模式；最后，采用迭代学习方法来捕捉嵌入表示和虚拟邻居预测之间的交互，对规则进行更新学习。图 12 给出了 VN 网络的框架。框架主要包括三个部分：(1) 基于规则的虚拟邻居预测；(2) 基于加权图卷积网络(Weighted Graph

chinaXiv:202305.00174v1

Convolutional Network, WGCN)[87]捕获结构信息并嵌入实体的编码器；(3)计算边缘概率并使用软标签对模型进行细化的解码器。VN 网络不仅推理精度更高，对于邻居稀疏问题也具有显著更强的鲁棒性。

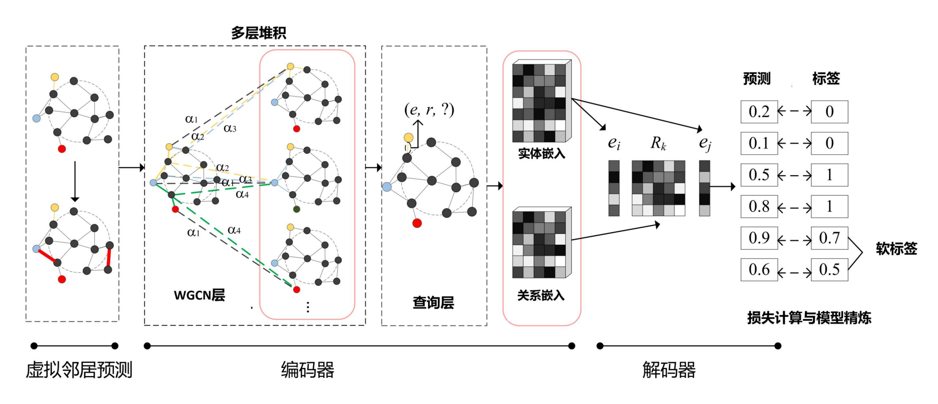


图 12　VN 网络框架

MLN 能够利用一阶逻辑的领域知识，同时处理不确定性，但难以适应复杂图结构，推理困难。神经网络建模有较高的推理效率，但无法利用领域知识。Qu 等[89]提出了概率逻辑神经网络 pLogicNet，它结合了 MLN 和神经网络嵌入表示学习的优点。pLogicNet 采用具有一阶逻辑的马尔可夫逻辑网络定义所有可能的三元组的联合分布，并利用变分 EM 算法对其进行高效优化。在 E 步中，将变分分布参数化为知识图谱嵌入模型，并使用平摊平均场(Amortized Mean-Field)[90]推断来计算未观测三元组的似然概率，以推理未观测到的三元组。在 M 步中，根据观察到的三元组和预测的三元组计算伪似然概率，以更新逻辑规则的权重，最后使用梯度下降法进行训练。

类似的，Vardhan 等[92]提出了基于 MLN 和图注意力网络[91]的概率逻辑图注意力网络(probabilistic Graph Attention neTwork, GAT, pGAT)。pGAT 采用变分 EM 算法优化 MLN 定义的所有可能三元组的联合分布，进而有效地结合一阶逻辑和图注意网络。在 E 步骤中，使用 GAT 嵌入来推断未观察到的三元组；在 M 步中，根据观测到的三元组和从上述嵌入中得到的推断三元组对马尔可夫逻辑网络规则的权值进行更新。pGAT 方法的框架如图 13 所示。pGAT 将马尔可夫逻辑网络与图神经网网络相结合，既利用了 MLN 的概率逻辑规则表达能力，也利用了图注意力网络能够捕获相邻实体特征、适应高度复杂图问题的能力，能够应用于大规模的复杂知识图谱推理。

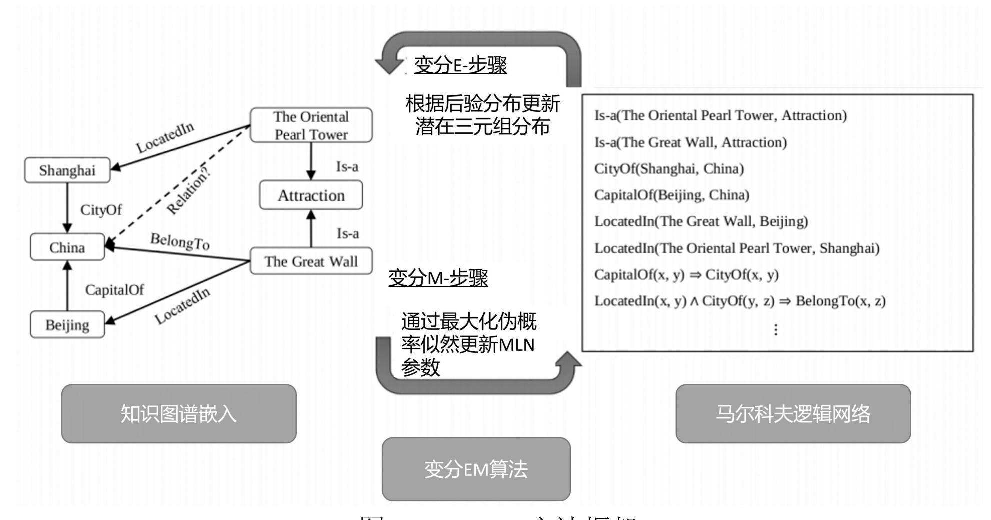


图 13　pGAT 方法框架

chinaXiv:202305.00174v1

chinaXiv:202305.00174v1

此外，Zhang 等[93]也对如何结合图神经网络(Graph Neural Network, GNN)与 MLN 进行了研究探索，他们提出了 ExpressGNN 方法。ExpressGNN 的本质是 MLN 的推理网络，通过利用编码在逻辑规则中的先验知识，以及 GNN 捕获隐含编码在知识图中的结构知识，将 MLN 推理扩展到更大的知识图谱推理问题。图 14 给出了 ExpressGNN 的模型框架，首先，用 MLN 对所有观测变量和潜在变量的联合概率分布进行建模，通过优化数据对数似然的变分证据下界(Evidence Lower BOund, ELBO)求解；然后，在 E-步骤中采用平均场分布来近似真实后验分布，并使用 GNN 来参数化变分后验；最后，在 M-步骤中，只保留基础谓词来计算伪似然，学习 MLN 中变分后验固定的逻辑公式的权重。ExpressGNN 不仅有更高的推理精度，还能通过调整 GNN 和嵌入部分的维度来权衡模型的紧凑性和表达性。此外，ExpressGNN 可以处理目标谓词较少的少次学习(few-shot)或零标记实例的零次学习(zero-shot)问题。

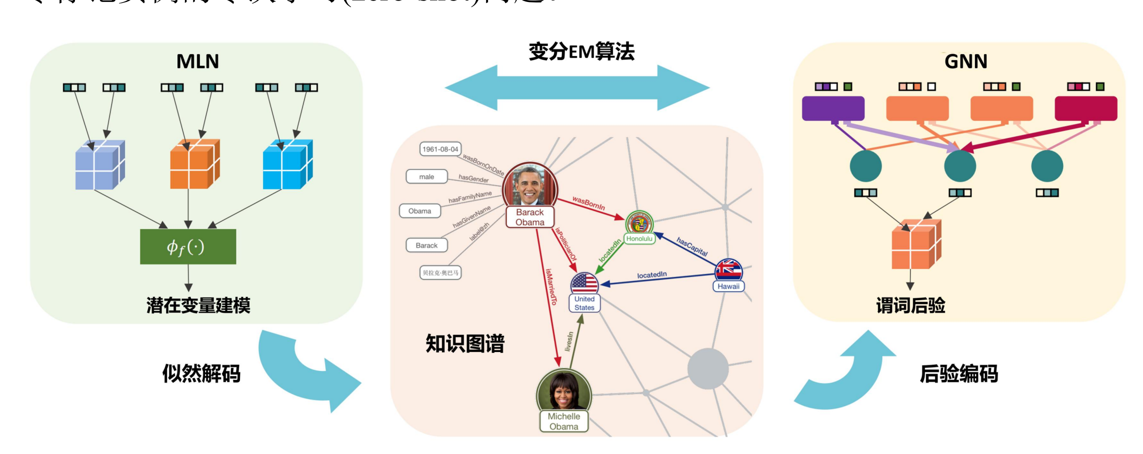


图 14　ExpressGNN 的模型框架

此外，还有部分学者将规则作为引导强化学习智能体路径探索的辅助知识，进行知识图谱推理。例如，Lei 等[94]提出了 RuleGuider 方法，首先使用 AnyBURL 方法生成的高质量规则为智能体提供奖励监督，然后基于规则及其实体变量的结构，将智能体分解为两个子智能体，即实体智能体和关系智能体。实体智能体将为可选择的实体生成实体分布；关系智能体对关系进行采样后，将根据所选关系修剪实体空间。实体代理然后基于修剪后的实体空间对实体进行采样。在最后一步，他们根据最后选择的实体获得命中奖励，并根据选择的关系路径从预先挖掘的规则集中获得规则指导奖励。

Xia 等[95]针对稀疏知识图谱的推理，提出了一种具有强化学习的混合多跳推理方法 SparKGR。该方法使用规则引导来动态地完成缺失路径，以增加强化学习智能体的动作空间，以缓解知识图谱的不完整性和稀疏性；然后设计了规则归纳和事实推理的迭代优化方法，以结合来自知识图谱的全局信息，来指导智能体的探索。SparKGR 能够有效降低知识图谱的稀疏性，提升路径探索效率，在保证可解释性的前提下有效提升推理准确度。智能体利用动态路径补全、迭代规则引导和策略函数引导 3 种策略选择下一步的动作空间，最后将 3 种策略的概率分布加权求和，确定智能体的下一步动作。图 15 呈现了 SparKGR 的动态路径补全策略和迭代规则引导策略。

利用规则改进神经网络建模的方法可以为神经网络的训练和参数学习等提供数据辅助和方向指导，使得推理结果具有一定的可解释性，但方法没有从模型内部结构和参数揭示规律，未能从根本上改变神经网络的“黑盒”性质。规则和

深度强化学习相结合的方法具有较强的可解释性，但其可解释性来自于智能体对路径的探索，而并非神经网络本身。

6.2 神经网络辅助规则学习的知识图谱推理

此类方法通过发挥神经网络的优势来处理数据的不确定性和模糊性问题，缩小规则推理的搜索空间，以发现更高质量的规则用于知识图谱推理，主要包括基于神经定理证明的方法、基于矩阵乘法的方法、基于神经归纳学习的方法等。

(1)基于神经定理证明的方法

基于神经定理证明的方法以 Prolog 编程语言[96]为基础。Prolog 是一种演绎推理逻辑编程语言，包含谓词匹配、规则求解、回溯判别等步骤，通过递归匹配算法与神经网络嵌入表示相结合，执行知识推理，并提供可解释的定理证明路径。

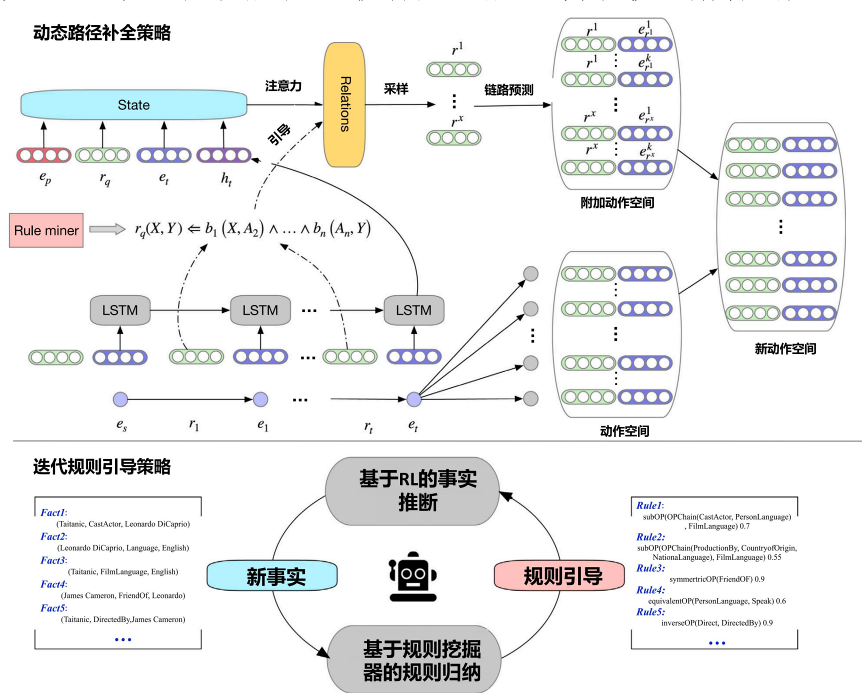


图 15 SparKGR 的动态路径补全策略和迭代规则引导策略

Rocktaschel 等[97]结合 Prolog 和 ILP，提出了端到端的知识推理方法 NTP(Neural Theorem Prover)。NTP 借鉴 Prolog 的反向链接算法递归构建神经网络，对符号的密集向量表示进行操作，然后利用可微证明器(Differentiable Prover, DP)对知识图谱的推理进行端到端可微证明。NTP 以原子、规则及其证明状态作为输入，并返回一个新的证明状态列表。首先利用 DP 的三个模块来证明知识图谱上的查询：(1)统一模块，用径向基函数核对规则原子的向量表示进行可微计算来代替符号统一，从而将规则推理与学习原子和规则的向量表示相结合；(2)OR 模块，在知识图谱中应用规则；(3)AND 模块，用于验证成功或失败，并与 OR 模块递归，用于证明状态和子目标。然后，通过证明汇总来计算目标的证明成功分数。最后，利用梯度下降算法将 NTP 用于 ILP，替代传统的组合搜索策略，对神经网络进行训练，并借助预测误差反向传播从数据中学习可解释的规则。NTP 可以从给定的不完整知识图谱中推断事实，还引入了可解释的一阶逻辑规则。然而，该方法的执行过程需要枚举和评价给定目标的所有有界深度证明路径，这导致 NTP 方法在大型知识图谱中效率低下。

针对 NTP 难以适用于大型知识图谱的问题，Minervini 等[98]提出了其改进方法 GNTP(Greedy NTP)，通过在共享的嵌入空间中嵌入逻辑事实和自然语言句

chinaXiv:202305.00174v1

子，对知识库和文本进行联合推理。针对 NTP 存在的问题，GNTP 利用三种技术予以处理：(1)使用事实嵌入来选择用于证明子目标的最近邻事实，使用关系嵌入来选择要扩展的规则，从而只考虑与最大证明分数相关的证明路径的子集，减少候选证明路径的数量，降低空间复杂性；(2)引入规则归纳的注意力机制来处理已知谓词，提高模型的参数效率，降低时间复杂性；(3)将 NTP 扩展到自然语言，使用端到端可微读取部分，在共享空间中联合嵌入谓词和文本表面模式，以充分利用自然语言文本信息，为推理提供更多知识。GNTP 在保持 NTP 方法可解释性的基础上，缓解了其复杂性和可迁移性的限制，使其能够应用于真实世界的大规模知识图谱。

NTP 在训练过程中需要考虑解释给定目标或子目标的所有规则，导致其在大量规则或推理步骤的情况下效率显著降低。为了缩小规则搜索空间，Minervini 等[99]又提出了 CTP(Conditional Theorem Prover)方法。该方法通过构建神经网络动态生成最小规则集，并采用基于梯度的优化来学习最优规则选择策略。对于每个子目标，CTP 的选择模块首先动态生成最小规则集合，以便在后续的推理步骤中使用。然后，对规则选择提出了三种重组方法，即神经目标重组、注意力目标重组、基于记忆的目标重组。其中，神经目标重组将“选择”操作定义为目标谓词的线性函数，以子目标为条件生成一组较少的规则，还可以在下游推理任务中进行端到端训练；注意力目标重组基于谓词生成关系集合上的注意力分布，以在选择模块中加入有用的先验；基于记忆的目标重组模块将规则存储在可微分存储器中，它可以在给定目标的前提下，在规则上产生注意力分布，以用于关键值记忆查找。相比 NTP，CTP 具有更强的可迁移性和更高的推理准确度。

此类方法的缺点是，进行定理证明往往需要进行计算量较大的规则搜索，但规则搜索的精度和效率难以权衡兼顾。此外，神经定理证明方法本身是一种演绎推理方法，无法发现新的规则。

(2)基于矩阵乘法的方法

此类方法将关系定义为二维矩阵，用矩阵乘法来表示多跳推理，并利用循环神经网络来模拟规则推理。矩阵乘法是可微的，因此模型可以实现规则参数和规则结构的同步学习，处理复杂的知识图谱推理任务。此外，推理过程是由每个层次的矩阵乘法逐层运算得到的，可以为推理结果提供更清晰的解释依据。

Cohen 提出了概率演绎数据框架 TensorLog[100]，该框架可以推断加权链式逻辑规则，以解释知识图谱中的关系，进而用于知识推理。在 TensorLog 中，知识图谱的每个实体用独热嵌入表示，每个关系 R 被表示为矩阵 M_R，其中如果 $R(i,j)$ 在图中，则每个元素 $M_R[i,j]=1$。然后，给出一个规则 γ：$R(X,Y)\Rightarrow P(X,Z)\wedge Q(Z,Y)$ 以及头实体 X，可以通过执行矩阵乘法 $\mathbf{M}_P\cdot\mathbf{M}_R\cdot\mathbf{v}_X$ 来进行答案的逻辑推理。由于查询关系可以由多条规则解释，因此通过组合所有规则来计算查询关系的得分，即：

$$\sum_{\gamma}\alpha_{\gamma}\prod_{k\in\beta_{\gamma}}\mathbf{M}_{R_k}, \tag{16}$$

其中 γ 表示所有可能的规则，α_{γ} 是与 γ 相关的置信度，β_{γ} 是 γ 中所有谓词的有序列表。在推理过程中，给定头部实体 X，每个正确答案 Y 满足 $R(X,Y)$ 的概率被最大化，即：

chinaXiv:202305.00174v1

$$\max_{\{\alpha_\gamma,\beta_\gamma\}} \sum_{X,Y} \mathbf{v}_Y^T \left(\sum_\gamma \alpha_\gamma \prod_{k\in\beta_\gamma} \mathbf{M}_{R_k} \mathbf{v}_X \right), \tag{17}$$

TensorLog 首先将逻辑规则中的每个子句转换为因子图，图中的每个节点表示规则中的一个变量，每条边表示规则中一个谓词或关系。然后，对因子图的不同类型查询，执行信念传播(Belief Propagation, BP)，所需的消息传递步骤被“展开”为一个可微函数，进而计算概率。最后将这些可微函数进行递归组合，可以实现在包含多个相互关联的规则和谓词的非平凡逻辑理论中执行推理。

TensorLog 以可微矩阵生成的方式估计规则置信度，但是无法生成新的逻辑规则。此外，由于每个规则都与一个参数相关联，而枚举规则的计算复杂度很高，因此参数学习十分困难。针对 TensorLog 的缺点，Yang 等[101]提出了神经逻辑编程(Neural Logic Programing, Neural LP)，对 TensorLog 的矩阵推理框架进行了改进。Neural LP 交换了对式 17 中的总和及乘积，并将规则的权重分解为规则中谓词的权重。为了处理不同长度的规则，Neural LP 还设计了一种递归方法来对规则进行动态建模：

$$\mathbf{u}_0 = \mathbf{v}_X,$$

$$\mathbf{u}_t = \sum_{k}^{|\mathbf{R}|} a_t^k \mathbf{M}_{R_k} \left(\sum_{\tau=0}^{t-1} b_t^\tau \mathbf{u}_\tau \right) \text{ for } 1 \le t \le T, \tag{18}$$

$$\mathbf{u}_{T+1} = \sum_{\tau=0}^{T} b_{T+1}^\tau \mathbf{u}_\tau,$$

T 是规则的最大长度，$|\mathbf{R}|$ 是知识图谱中的谓词数量，$\mathbf{u}$ 是输入实体 $\mathbf{v}_X$ 的辅助存储向量。在每一次递归中，首先使用记忆注意力向量 $\mathbf{b}_t$ 计算先前辅助存储向量的加权平均值，然后利用操作注意力向量 $\mathbf{a}_t$ 进行 TensorLog 运算。最后，计算所有辅助存储向量的权重平均值，并使用注意力来控制规则的长度。对使用 RNN 建模对 $\mathbf{a}_t$ 和 $\mathbf{b}_t$ 进行建模求解，并以其为依据从重建规则及其置信度。

Neural LP 在处理年龄、体重、科学测量等数字特征方面受到限制，导致其能够表示的规则类型受限。为此，Wang 等[102]提出了 Num-Neural LP 方法，以学习数值型规则。该方法在 Neural LP 的基础上，通过使用动态规划和累积和运算来表达数值比较器运算符，并将其与稀疏运算符联合使用，以适应密集的运算。此外，Num-Neural LP 还添加了两种运算符：一是分类运算符，用于比较实体属性上的函数；二是否定运算符，基于局部封闭世界假设，对于给定的取反算子，只翻转至少包含一个非零元素行中的元素。Num-Neural LP 能够聚合提取表达能力更强的规则，学习到的规则具有更高的质量，推理结果也更准确。

为从知识图谱中挖掘变长一阶逻辑规则，Sadeghian 等[103]提出了端到端可微模型 DRUM。DRUM 能同时学习规则结构和置信度分数，实现归纳和可解释的知识图谱推理。首先，将规则推理转换为向量之间的连乘计算。然后，引入邻接矩阵为单位阵 $\mathbf{I}$ 的特殊关系 B_0，以学习变长的规则，并引入置信度张量来避免学习到置信度较高的错误规则。最后，利用 Bi-RNN 能够捕捉原子在规则中的前后顺序信息的特点，建模规则头和规则体中关系的关联。除了推理准确度方面的优势外，DRUM 还可以进行归纳式推理，处理未出现过的实体，具有较强的可解释性。

受限于矩阵乘法的性质，TensorLog、Neural LP 和 Num-Neural LP 等多数方法只能学习链式逻辑规则，无法适用于树状和合取等复杂的规则形式。基于特定

chinaXiv:202305.00174v1

的头实体来推断规则会影响此类方法的泛化能力。此外，由于该类方法中，实体、关系和规则到矩阵的转化依赖于嵌入表示方法，因此模型本身的鲁棒性不强。

(3)基于神经归纳学习的方法

此类方法利用图神经网络、循环神经网络等进行归纳式逻辑规则学习，进而执行推理预测。图神经网络以图数据作为表征方式，能够对节点和关系进行优化计算，并且考虑了节点的隐层信息、邻居节点信息和图结构信息，非常适用于知识图谱的规则学习。而循环神经网络在建模具有顺序信息的数据中表现出色，同样能够应用于知识图谱推理中。神经归纳学习在近几年成为了学者们研究的重点内容。

Yang 等[104]提出了神经逻辑归纳学习(Neural Logic Inductive Learning, NLIL)方法。NLIL 通过结合原始语句(Primitive Statement)来处理非链式规则。原始语句是指适用于逻辑变量或某些运算符结果的谓词。一元原始语句包含从某个变量开始的关系路径，二元原始语句包含分别从两个不同变量开始的两条关系路径，原始语句还能通过$\{\wedge,\vee,\neg\}$运算符进行逻辑组合，这使其具备表示树状和合取逻辑规则的能力，如图 15 所示。在规则学习中，为了确定不同原始语句逻辑组合的重要性权重，利用三个堆叠的 Transformer[105]网络来学习不同的注意力向量。

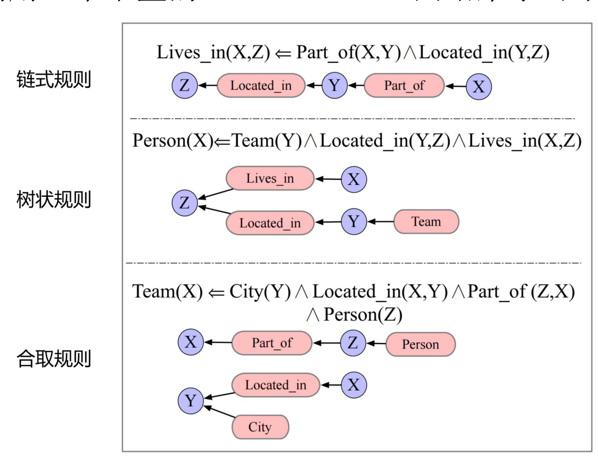


图 16　NLIL 能够学习的规则类型

Dong 等[106]提出了一种用于归纳学习和逻辑推理的神经符号体系结构——神经逻辑机(Neural Logic Machines, NLM)。NLM 利用神经网络作为函数逼近器，运用逻辑编程处理具有性质、关系、逻辑连接词和量词的对象。在内部，NLM 通过在一组固定的对象上将谓词固定为 True 或 False，进而使用张量来表示逻辑谓词。然后，利用多层感知机(Multi-Layer Perceptron, MLP)建立神经体系结构来学习规则，该体系结构中使用布尔逻辑和扩展、缩减等量化手段来表示元规则。在对小规模任务进行训练后，NLM 可以恢复提升的规则，并推广到大规模任务。

Chen 等[107]提出了一种基于 GNN 的图协同推理(Graph Collaborative Reasoning, GCR)方法，从逻辑推理的角度使用相邻链接信息对知识图谱进行关系推理，并利用 GNN 的消息传递功能来聚合相邻链路的丰富信息，以实现高效推理。首先把图结构转换为 Horn 规则的逻辑表达式，从而将知识推理任务转换为神经逻辑推理问题。然后将每个三元组编码为谓词嵌入，并将每个关系建模为神经模块来编码三元组。最后，通过编码的谓词嵌入，根据建模的 Horn 规则使用神经模块来构建网络结构，并使用反向传播来学习模型参数。GCR 可以充分利用链接之间的逻辑关系，不需要手动预定义规则，具有较强的可解释性和可迁移性，特别是适用于数据稀疏的知识图谱推理。

chinaXiv:202305.00174v1

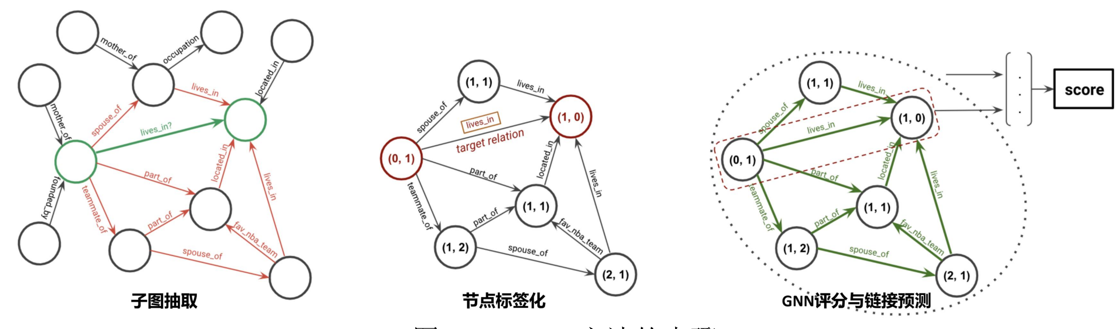


图 17　GraIL 方法的步骤

为在含有不完备训练集的知识图谱中实现高效的归纳式推理，Teru 等[108]提出了一种基于图神经网络的关系预测框架 GraIL(Graph Inductive Learning)。GraIL 可以挖掘一阶逻辑规则，通过学习子图嵌入表示，从候选关系周围的子图结构预测关系。GraIL 主要包括三个步骤：(1)子图提取，提取目标节点周围的封闭子图，以挖掘包含推导目标节点之间关系所需的逻辑证据，对规则进行显式编码；(2)节点标签化，利用节点与周围子图节点的距离构建节点特征矩阵，作为 GNN 的输入，以初始化神经消息传递算法；(3)GNN 评分，使用 GNN 对三元组的可能性进行评分。使用神经消息算法，通过将节点表示与邻居表示的聚合（aggregate）相结合（combine）来迭代更新节点表示。例如，对于 GNN 的第 k 层，节点表示的迭代更新公式如下：

$$\begin{gathered}\mathbf{a}_t^k=\mathrm{AGGREGATE}^k\left(\left\{\mathbf{h}_s^{k-1}:s\in\mathcal{N}(t),\mathbf{h}_t^{k-1}\right\}\right),\\ \mathbf{h}_t^k=\mathrm{COMBINE}^k\left(\mathbf{h}_t^{k-1},\mathbf{a}_t^k\right),\end{gathered}\tag{19}$$

$\mathbf{a}_t^k$ 是来自邻居的聚合消息，$\mathbf{h}_t^k$ 表示第 k 层中节点 t 的潜在表示，$\mathcal{N}(t)$ 表示节点 t 的直接邻居集。AGGREGATE 函数基于多关系 R-GCN[126]和边缘注意力定义，COMBINE 函数由 R-GCN 的 ReLU 函数确定。最后，利用子图表示 $\mathbf{h}^L_{\mathcal{G}(u,v,r_t)}$、目标节点的潜在表示 $\mathbf{h}_u^L,\mathbf{h}_v^L$ 和目标关系的嵌入表示 $\mathbf{e}_{r_t}$ 来计算三元组 (u,r_t,v) 的可能性分数，即：

$$score(u,r_t,v)=\mathbf{W}^T\left[\mathbf{h}^L_{\mathcal{G}(u,v,r_t)}\oplus\mathbf{h}_u^L\oplus\mathbf{h}_v^L\oplus\mathbf{e}_{r_t}\right],\tag{20}$$

根据可能性分数的大小进行推理。图 17 展示了 GraIL 方法的流程步骤。GraIL 可以学习一阶逻辑规则的有用子集，推理出训练集中不存在的实体，相比其他归纳式推理方法有着更强的精度和可解释性。

GraIL 方法开启了利用图神经网络进行归纳式逻辑规则学习的研究序幕。之后，研究人员开展了多项相关改进工作。针对 GraIL 在提取目标三元组的封闭子图时忽略了知识图谱的方向性质，导致其无法处理不对称和反对称关系的问题，Mai 等[109]基于 GraIL 提出了 CoMPILE 方法，利用节点-边通信消息传递机制来取代原来的 GNN，以衡量关系的重要性。该方法首先为三元组提取有向封闭子图，然后扩展通信消息传递网络框架，以加强实体和关系之间的信息交互，同时更新边缘和实体嵌入。最后，利用边感知注意力机制来聚合局部邻域特征，并收集全局实体信息，以丰富实体和关系表示。相比 GraIL，CoMPILE 能够在控制模型参数数量的前提下，在有向子图中执行通信消息传递，以处理不对称和反对称关系。

chinaXiv:202305.00174v1

chinaXiv:202305.00174v1

此外，还有一些归纳式知识逻辑规则学习方法，不是以 GraIL 为基础改进的，但相比 GraIL 在性能上都有所提升。例如，针对 GraIL 方法没有考虑语义相关性的问题，Chen 等[110]提出了 TACT(Topology-Aware CorrelaTions)方法，考虑子图中关系之间的相关性，对关系相关网络进行编码，以增强封闭子图的编码，执行图的拓扑感知校正。针对实体独立关系建模需求和离散逻辑推理可解释性难以权衡的问题，Lin 等[111]提出了 ConGLR 方法，将上下文图和逻辑推理结合起来，在抽取目标头尾实体的封闭子图的基础上，构建包含关系路径、关系和实体的上下文图，并利用感知注意力的图卷积网络处理子图和上下文图，最后通过将关系路径作为规则体，目标关系作为规则头，结合神经计算和逻辑推理，计算置信分数。针对稀疏子图难以处理的问题，Xu 等[112]提出了 SNRI(Subgraph Neighboring Relations Infomax)方法，利用相邻关系特征和相邻关系路径来捕捉子图中实体的完全相邻关系，并通过互信息最大化对相邻关系路径进行全局建模，进而将完整的关系信息合并到封闭子图中，建模相邻的关系路径，提高归纳推理预测的性能。Zha 等[113]提出的 BERTRL(BERT-based Relational Learning)将预训练语言模型与关系学习相结合，首先利用对知识图进行线性化处理，作为 BERT 的有效输入，然后建立 BERT 模型对推理路径进行编码，最后利用 bag scoring 方法聚合三元组路径和推理路径的评分，BERTRL 的框架如图 18 所示。

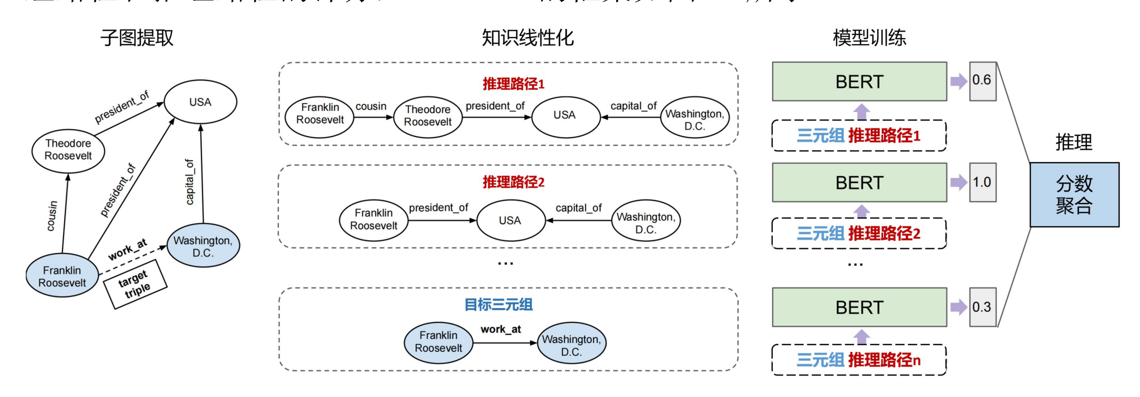


图 18　BERTRL 的模型框架

为了从潜在规则中获得实体相关的关系语义，进而解决子图中规则稀缺导致的监督不足问题，Pan 等[114]基于图卷积网络提出了归纳路径的关系对比方法 RPC-IR(Relational Path Contrast for Inductive Reasoning)方法。RPC-IR 包含三个步骤：(1)从目标三元组的封闭子图中提取路径，并产生关系路径的正负样本进行对比；(2)使用 GCN 获得正样本和负样本的嵌入表示；　(3)用子图和关系路径对目标三元组进行评分，并使用一种同时考虑监督信息和对比信息的策略进行联合训练。通过这些步骤，根据关系路径的置信度确定规则。

GraIL 及其改进方法将规则视为路径，而大规模图谱中的路径挖掘往往面临巨大的搜索空间。针对该问题，Yan 等[115]提出了 CBGNN(Cycle Basis Graph Neural Network)方法，从代数拓扑的角度将逻辑规则视为循环，建立 GNN 模型，利用其消息传递功能在循环空间中运行隐式代数运算，以学习循环的表示，然后通过探索循环空间的线性结构学习规则，进而执行知识推理。图 19 给出了 CBGNN 的模型结构。方法主要包含两个阶段。在第一阶段，构建循环基，并为每个循环基构建循环图。在循环图中，节点表示基中的循环，如果节点对应的循环具有较强的交互作用，则认为节点是连接的。循环的信息可以转换为第二阶段的新图中的节点特征。第二阶段，在循环图上建立 GNN，以学习循环的置信值，并将其映射为目标三元组的置信值。在该阶段中，不同的循环基构建不同的 GNN，

但这些 GNN 共享权重，并可以通过对其进行聚合计算三元组的置信度。相比基于规则路径的方法，CBGNN 方法具有更高的效率，同时为将高级拓扑信息纳入图表示学习提供了启发。

针对神经逻辑编程方法搜索空间过大和强化学习方法奖励稀疏导致的推理效率问题，Qu 等[116]提出了 RNNLogic 模型。RNNLogic 将逻辑规则视为一个潜在变量，并用逻辑规则训练规则生成器和推理预测器。其中，规则生成器使用 RNN 参数化。在模型优化的过程中，采用一种基于 EM 的优化算法。在每次迭代中，首先对推理预测器进行更新，以探索用于逻辑推理规则。然后，在 E 步骤中，通过后验推理，从所有生成的规则中选择一组高质量的规则；在 M 步骤

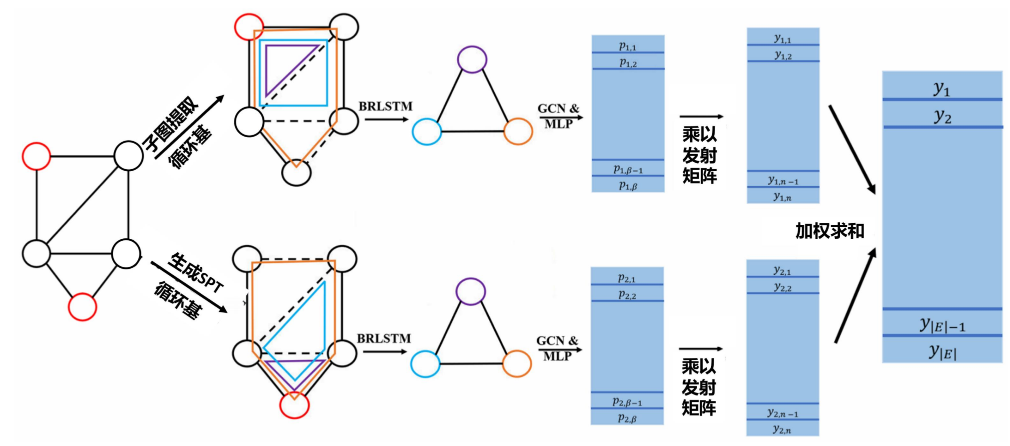


图 19　CBGNN 的模型结构

中，使用在 E 步骤中选择的规则来更新规则生成器。RNNLogic 的模型框架如图 20 所示。其中，q 代表查询，$\mathbf{z}$ 表示规则的潜在表示，θ 和 w 表示模型参数，$p_\theta(\mathbf{z}|\mathbf{q})$ 表示由查询生成规则潜在分布的先验概率，$p_w(\mathbf{a}|\mathcal{G},\mathbf{q},\mathbf{z})$ 表示根据知识图谱、规则潜在表示和查询确定的答案的概率似然。

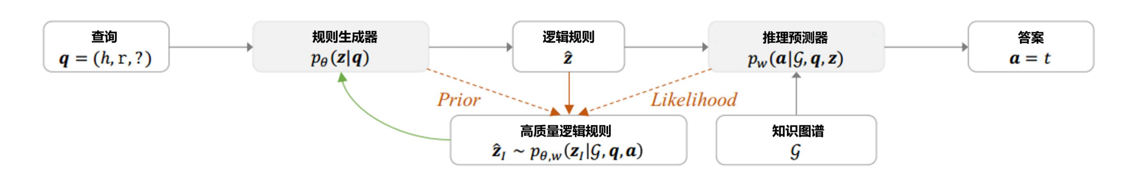


图 20　RNNLogic 的模型结构

此外，针对小样本(few-shot)知识图谱推理问题，Huang 等[117]提出了连接子图推理器方法(Connection Subgraph Reasoner, CSR)，将小样本知识推理转化为归纳推理问题。该方法不需要预训练环节，可以直接对目标少镜头任务进行预测。CSR 设计了一种新的基于图神经网络的编码器-解码器架构，来对支持和查询三元组之间的共享连接子图进行建模，并设计了自监督预训练方案以重建自动采样的连接子图。

针对知识图谱推理中的一次学习(one-shot)问题，Du 等[118]提出了 CogKR(Cognitive Knowledge Graph Reasoning)方法，该方法包含摘要和推理两个模块。摘要模块总结给定实例的基本关系，推理模块则根据摘要模块总结的关系进行推理。在摘要模块中，利用 GNN 生成的实体对嵌入表示获取实体对的关系。在推理模块中，CogKR 借鉴认知科学中的双过程理论，建立了两个系统，用于构建认知图以存储检索到的信息和推理结果。系统 1 为迭代协调检索系统，可以直观地收集推理相关证据，系统 2 为推理系统，对收集的信息进行关系推理。认

chinaXiv:202305.00174v1

知图的结构信息使模型能够聚合来自多个推理路径的证据，还能以图的形式给出推理过程的解释。最后，使用强化学习方法将图形结构建模转化为策略优化问题，并利用 REINFORCE 算法[124]进行策略优化。

总的来说，此类方法的效率和精度都比较高，也具备一定的可解释性。但由于利用神经网络学习规则，也在一定程度上引入了不可解释因素。表 6 给出了基于神经网络和规则的推理方法的相关信息。

7 比较与分析

从基于归纳逻辑编程的推理方法、基于概率图模型和规则的推理方法、基于嵌入表示和规则的推理方法到基于神经网络和规则的推理方法的发展，体现了基于规则的知识图谱推理方法的不断改进和完善。总的来看，四类方法的性能是递进提升的，但是效率、可解释性和可迁移性等性能方面，不同类别的方法及其子方法各有优缺点。

推理精度。知识图谱推理的主要评价指标包括平均倒数秩(Mean Reciprocal Rank, MRR)、前 k 次命中正确预测结果的比例 Hit@k(k=1,3,10)和准确度召回率曲线 AUC-PR 等，式 21 和式 22 给出了指标 MRR 和 Hit@k 的计算公式:

$$MRR=\frac{1}{2|\mathcal{T}_{test}|}\sum_{i\in\mathcal{T}_{test}}\left(\frac{1}{rank_i^s}+\frac{1}{rank_i^o}\right), \tag{21}$$

$$Hit@k=\frac{1}{2|\mathcal{T}_{test}|}\sum_{i\in\mathcal{T}_{test}}\left\{I\left[rank_i^s\le k\right]+I\left[rank_i^o\le k\right]\right\}, \tag{22}$$

其中 $\mathcal{T}_{test}$ 表示测试集事实数量，$rank_i^s$ 和 $rank_i^o$ 分别表示头实体推理和尾实体推理的排名，$I\left[P\right]$ 为指示函数，条件 P 成立则返回 1，否则返回 0。综合参考文献给出的实验结果，汇总了本文介绍的方法在经典的知识图谱推理数据集 FB15K-237[121]上，执行链路预测任务的实验结果，如表 7 所示。其中，未使用相关数据集、未采用上述指标、在实验中对数据集进行了划分处理或只针对某项任务(如实体预测)的，表格将其留空，只在 FB15K 数据集(未去掉反向关系)进行过实验的，在对应指标下方用横线标记，选取了数据集的 few-shot 子集的，在对应指标下方用波浪线标记。早期的部分方法由于没有在基准数据集上实验，后续研究中也较难复现，因此未能给出实验数据。从表 7 中不难看出，基于神经网络和规则的知识图谱推理的准确度总体高于其他三类方法。

推理效率。基于归纳逻辑编程的推理方法中，基于一阶归纳学习的方法的计算复杂度随着规则数量指数上升，在中大规模的知识图谱中效率较低；基于关联规则挖掘的方法由于需要规则遍历，难适用于在大型知识图谱；基于关系路径采样的方法相比前两类方法，效率有明显提升，可以用于在大规模知识图谱中学习规则以用于推理。基于概率图模型和规则的推理方法，由于涉及复杂的搜索和概率计算，导致其整体效率都比较低，大多难以应用于大规模知识图谱。基于嵌入表示和规则的推理方法使用表示学习方法缓解了原有的搜索、遍历、采样等流程带来的耗时，因此比前两类方法的推理效率更高，其中，由于占主导地位的技术不同，规则嵌入表示学习的方法效率高于嵌入表示与规则相互增强的方法。最后，基于神经网络和规则的推理方法效率显著优于前三类方法。

chinaXiv:202305.00174v1

表 6 基于神经网络和规则的推理方法的相关信息

分类		特点	经典方法	年份	代码地址	不足
规则改进神经网络建模的方法		聚焦利用逻辑规则提升神经网络嵌入和预测能力。利用原始观察到的三元组和一些预定义规则推断的三元组或规则。规则指导神经网络训练，添加的三元组作为增广数据集，提升神经网络可解释性	LAN	2019	https://github.com/wangpf3/LAN	没有从模型内部结构和参数揭示规律，未能改变神经网络的“黑盒”性质
			GRAFT-Nets	2019	https://github.com/OceanskySun/GraftNet	
			VN	2020	https://github.com/Yongquan-He/VN-Network	
			pLogicNet	2019	https://github.com/DeepGraphLearning/pLogicNet	
			pGAT	2020	https://github.com/Harsha2399/pGAT	
			ExpressGNN	2020	https://github.com/expressGNN/ExpressGNN	
神经网络辅助规则学习的方法	基于定理证明	以演绎推理逻辑语言 Prolog 为基础，利用递归匹配算法与神经网络执行知识推理，提供可解释路径	NTP	2017	/	规则搜索精度和效率难兼顾，是一种演绎推理方法，无法发现新规则
			GNTP	2019	https://github.com/uclnlp/gntp	
			CTP	2020	https://github.com/uclnlp/ctp	
	基于矩阵乘法	将关系定义为二维矩阵，用矩阵乘法来表示多跳推理，并利用循环神经网络来模拟规则推理，可以实现规则参数和规则结构的同步学习，推理过程更清晰,可解释性更强	TensorLog	2016	https://github.com/TeamCohen/TensorLog	矩阵化依赖于嵌入表示方法，鲁棒性不强，对非链式规则的学习能力不强
			Neural LP	2017	https://github.com/fanyangxyz/Neural-LP	
			DRUM	2019	https://github.com/alisadeghian/DRUM	
			Num-Neural	2020	/	
	基于神经归纳学习	利用图神经网络、循环神经网络进行归纳式逻辑规则学习，进而执行推理预测。图神经网络能够对节点和关系进行优化计算，且考虑了节点的隐层信息、邻居节点信息和图结构信息，适用于知识图谱的规则学习。循环神经网络在建模具有顺序信息的数据中表现出色，适用于知识图谱推理	NLIL	2020	https://github.com/gblackout/NLIL	利用神经网络学习规则，引入了一定的不可解释因素
			NLM	2019	https://sites.google.com/view/neural-logic-machines	
			GCR	2021	/	
			GraIL	2020	https://github.com/kkteru/grail	
			CoMPILE	2021	https://github.com/TmacMai/CoMPILE_Inductive_Knowledge_Graph	
			TACT	2021	https://github.com/MIRALab-USTC/KG-TACT	
			ConGLR	2022	/	
			SNRI	2022	https://github.com/Tebmer/SNRI	
			BERTRL	2022	https://github.com/zhw12/BERTRL	
			CBGNN	2022	/	
			RPC-IR	2021	/	
			RNNLogic	2021	https://github.com/DeepGraphLearning/RNNLogic	
			CSR	2022	https://github.com/snap-stanford/csr	
			CogKR	2020	https://github.com/THUDM/CogKR	

chinaXiv:202305.00174v1

chinaXiv:202305.00174v1

表 7 基于规则的知识图谱推理方法在 FB15k-237 数据集上的链路预测准确度比较

类别	方法	MRR	Hit@1	Hit@3	Hit@10	AUC-PR
基于归纳逻辑编程	AMIE+	0.292	0.174		0.409	
	RuleN	/	0.182		0.420	
	C-NN	0.296	0.222		0.446	
	AnyBURL	0.31	0.233			
	SAFRAN	0.389	0.298		0.537	
	Rein-AnyBURL	0.354	0.267		0.527	
基于概率图模型和规则	MLN	0.098	0.067	0.103	0.16	
基于嵌入表示和规则	KALE	0.523	0.383	0.616	0.762	
	RUGE	0.768	0.703	0.815	0.865	
	TARE	0.781	0.721	0.839	0.902	
	C-NNE-AER	0.803	0.761	0.831	0.874	
基于神经网络和规则	LAN	0.394	0.302	0.446	0.566	
	VN	0.463	0.345	0.526	0.701	
	pLogicNet	0.332	0.237	0.367	0.524	
	pGAT	0.457	0.377	0.494	0.609	
	Neural LP	0.24				
	Num-Neural LP	0.259				
	NLIL	0.25				
	DRUM	0.343	0.255		0.516	
	GCR	**0.492**	**0.490**			
	GraIL				0.641	0.847
	CoMPILE				0.676	0.855
	TACT				0.741	0.830
	ConGLR				0.683	0.857
	SNRI				0.718	0.867
	BERTRL				0.620	0.695
	RPC-IR				0.676	0.872
	CBGNN				0.975	0.963
	RNNLogic	0.349	0.258	0.385	0.533	
	CSR	0.624	0.479		0.786	

可解释性。从可解释性角度评价，基于归纳逻辑编程的推理方法的可解释性最强，基于概率图模型和规则的推理方法由于涉及概率估计和模糊推理，其可解释性相对次之。基于嵌入表示和规则的推理方法以及基于神经网络和规则的推理方法，引入了不可解释的“黑盒”模型，对可解释性有一定的影响，这两类方法中，依赖于规则学习进行推理的方法的可解释性高于依赖于神经网络建模或嵌入表示的推理方法。

可迁移性。可迁移性是指模型或方法运用已有知识来学习新知识的能力。知识图谱推理方法的可迁移性，本文将其理解为方法在运用于不同知识图谱时的额外消耗和性能差别。总的看，受益于深度网络的灵活性和适应性优势，基于神经网络和规则的推理方法的可迁移性最强，其次是基于嵌入表示和规则的推理方法。基于归纳逻辑编程的推理方法、基于概率图模型和规则的推理方法的可迁移性相对较差，因为这两类方法面对不同的知识图谱需要重新尽可能地学习更多规则。

8 挑战与展望

通过与嵌入表示、神经网络等方法模型相结合，基于规则的知识图谱推理方法已逐渐改善了过往推理速度慢、计算复杂度高、表达能力有限等问题，在大规模知识图谱的推理中展现出良好的性能。然而，尽管基于规则的方法具有相对较强的可解释性、可靠性和泛化能力，其推理的精度、效率相比部分最先进的、单纯基于神经网络的推理方法仍有差距，在未来研究中，此类推理方法仍然面临许多新挑战，需要在未来的研究中加以解决，主要包括以下 7 个方面：

规则学习效率。从大规模知识图谱中学习规则是困难的。虽然研究者们通过结合机器学习、深度学习、强化学习等技术方法，已经大大改进了规则学习的效率，但目前绝大多数规则学习方法依赖于知识图谱的路径搜索。而路径的数量随着实体和关系的数量的增加，呈现指数上升的趋势。因此，需要研究如何采取有效的抽样策略减少路径搜索量，或者建立预训练模型，通过融合先验信息、文本信息、语料库来降低搜索复杂度。同时，如何将规则表示为更容易被神经网络接受的形式，进而提升规则学习和推理效果，也需要进一步探索。

规则质量评估。早期的规则质量评估，主要通过在封闭世界假设下计算置信度、支持度等指标实现，但这些指标在数据缺失的情况下难以计算。之后，Galárraga 等[47]提出了部分完整性（PCA）等假设，通过放松置信度计算公式中对负实例中的要求来缓解数据缺失问题，但该假设也只是折中方案。因此，如何设计更加合理、可行的规则评价指标，需要在下一步的工作中加以关注。另一方面，规则学习的目的是为了推理，规则的质量如何影响知识图谱推理的精度这一问题，也值得进行系统深入的研究，因为有时错误的规则也可能推理出正确的答案。

可解释性。基于规则的知识图谱推理方法的最大优势在于其可解释性，这赋予了其在实际应用中的安全性、可靠性。以规则表示为辅助，嵌入表示或神经建模主导的推理方法，在效率方面优于利用神经网络和嵌入表示辅助规则学习的推理方法，但不具备可解释性。神经网络和嵌入表示辅助规则学习的方法将规则转换到向量空间中进行推理，但这些向量或矩阵的中间表示只有被正确解释，才能达成真正的“模型可解释”。未来可以尝试将推理步骤嵌入模型以挖掘理解规则的中间表示。另一方面，可以利用可解释的机器学习方法集成规则和嵌入表示，以在保证可解释性的同时尽量提升效率。

基准数据集。当前知识图谱推理领域常用的数据集，如 FB15K-237、NELL-995[122]、WN18R[123]等，都是通过划分训练集、测试集、验证集，以评价比较有监督知识图谱推理方法的性能。当前，多数基于规则的知识图谱推理方法的性能好坏，也是在上述数据集上进行评判的。但这些数据集无法完全准确地表达和体现所有的内在规则和逻辑模式，有些错误或者冗余的规则也可能推出正确的结果，这对结果可解释性会产生误导式的影响，若将其应用在一些重要领域中，

chinaXiv:202305.00174v1

势必会导致严重后果。因此，未来应提出包含不同预定义规则或逻辑模式的基准数据集，以同时满足规则学习和知识图谱推理性能的验证。

时序知识图谱和多模态知识图谱推理。目前，大多数基于规则的推理方法研究都局限于静态知识图谱中，只有少数工作，如 Liu 等[119]提出的 Tlogic 研究了如何利用规则在时序知识图谱中进行推理。时序知识图谱推理需要挖掘时间依赖规则或者动态规则，所以可以尝试将将 RNN、LSTM 等能够建模时间信息的模型与现有的规则学习方法相融合，以学习到时序知识图谱的规则。多模态知识图谱包含文本描述、时序特征、图像信息等多模态信息，其规则的组成更复杂，表示及学习难度更大。因此，下一步要研究利用现有的多模态学习模型，通过数据融合、统一表示等方式，实现多模态知识图谱自动规则挖掘。

小样本和零样本知识图谱规则发现。对于 few-shot、one-shot 和 zero-shot 的情景，规则学习更加困难，特别是高阶的规则更加难以学习，这对推理的精度会产生较大影响。未来主要有两种解决思路，一是过融合附加信息增强少样本知识的表示能力，从中学习更高质量的规则；二是以现有的高质量知识图谱为先验，利用其高质量样本来辅助规则学习和推理。

将规则融入大语言模型。LLM 在运用于智能问答、推荐系统等知识推理领域时，容易受到信息冲突、信息混淆等因素的影响，导致其在面对长尾场景时可靠性不够稳定[120]，存在事实欺骗性风险。根本原因在于其在任务中“无意识”地利用知识，缺乏正确合理的引导。此外，在领域知识的问答和推理中，LLM 也面临着因数据稀疏而导致的“知识匮乏”问题。因此，如何将知识图谱和规则融入 LLM 之中，在训练前、训练中和训练后对预训练语言模型进行增强，也是未来研究中具有前景和潜力的一个方向。

chinaXiv:202305.00174v1

参考文献：

[1] 李志飞，赵月，张龑. 基于表示学习的知识图谱推理研究综述. 计算机科学, 2023, 50(3): 94-113. (Li Zhifei, Zhao Yue, Zhang Yan. Survey of Knowledge Graph Reasoning Based on Representation Learning. Computer Science, 2023, 50(3): 94-113.)

[2] George A. Miller. WordNet: A Lexical Database for English. Communications of the ACM, 1995, 38(11): 39-41.

[3] Lehann J, Isele R, Jakob M, Jentzsch A, Kontokostas D, Mendes PN, Hellmann S, Morsey M, van Kleef Patrick, Aues S, Bizer C. DBPedia-A large-scale, multiligual knowledge base extracted from wikipedia. Semantic Web, 2015, 6(2): 167-195.

[4] Fabian MS, Gjergji K, Gerhard Y. Yago: a core of semantic knowledge unifying wordnet and wikipedia. Proceedings of the 16th International World Wide Web Conference. Banff Alberta, Canada, 2007:697-706.

[5] Mitchell T, Cohen W, Hruschka E, et al.Never-Ending Learning [J]. Communications of the ACM, 2018, 61(5):103-115.

[6] Wu Wentao, Li Hongsong, Wang Haixun, et al. Probase: a probabilistic taxonomy for text understanding. Proceedings of the 2012 ACM SIGMOD International Conference on Management of Data. Scottsdale Arizona, USA, 2012:481-492.

[7] Xu Bo, Xu Yong, Liang Jiaqing, et al. CN-DBpedia: A Never-Ending Chinese Knowledge Extraction System. Proceedings of the International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems. Arras, France, 2017:428-438.

[8] Niu Xing, Sun Xinruo, Wang Haofeng, et al. Zhishi.me - Weaving Chinese Linking Open Data. Proceedings of the 10th International Semantic Web Conference. Arras, France, 2011:205-220.

chinaXiv:202305.00174v1

[9] Wang Xiang, He Xiangnan, Cao Yixin, et al. KGAT: Knowledge Graph Attention Network for Recommendation. Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. Anchorage AK, USA, 2019:950-958.
[10] He Sun, Zou Lei, Yu Xu Jeffrey,et al. Answering Natural Language Questions by Subgraph Matching over Knowledge Graphs. IEEE Transactions on Knowledge and Data Engineering, 2018, 30(5):824-837.
[11] Zhou Hao, Young Tom, Huang Minlie, et al. Commonsense Knowledge Aware Conversation Generation with Graph Attention.Proceedings of the 27th International Joint Conference on Artificial Intelligence, Stockholm, Sweden, 2018:4623–4629.
[12] Liang Ke, Meng Lingyuan, Liu Meng, et al. Reasoning over Different Types of Knowledge Graphs: Static, Temporal and Multi-Modal.arXiv:2212.05767v1.
[13] 官赛萍，靳小龙，贾岩涛，王元卓，程学旗. 面向知识图谱的知识推理研究进展. 软件学报, 2018, 29(10): 2966-2994.(Guan Saiping, Jin Xiaolong, Jia Yantao, et al. Knowledge Reasoning Over Knowledge Graph: A Survey. Journal of Software, 2018, 29(10): 2966-2994)
[14] Costas Mavromatis, George Karypis. ReaRev: Adaptive Reasoning for Question Answering over Knowledge Graphs. Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing, Abu Dhabi, The United Arab Emirates, 2022:2447-2458.
[15] Xian Yikun, Fu Zuohui, Muthukrishnan S,et al. Reinforcement Knowledge Graph Reasoning for Explainable Recommendation. arXiv:1906.05237.
[16] Tuan Yi-Lin, Sajjad Beygi, Maryam Fazel-Zarandi,et al. Towards Large-Scale Interpretable Knowledge Graph Reasoning for Dialogue Systems. arXiv:2203.10610.
[17] Mohamad Yaser Jaradeh, Kuldeep Singh, Markus Stocker, et al. Information extraction pipelines for knowledge graphs. Knowledge and Information Systems, 2023, 65:1989-2016.
[18] Kenneth Marino, Ruslan Salakhutdinov, Abhinav Gupta. The More You Know: Using Knowledge Graphs for Image Classification. Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), San Francisco, USA, 2017: 20-28.
[19] 马玉凤,向南,豆亚杰等. 军事系统工程中的知识图谱应用及研究. 系统工程与电子技术, 2022, 44(1):146-153.(Ma Yufeng, Xiang Nan, Dou Yajie, et al. Applications and research of knowledge graph in military system engineering. Systems Engineering and Electronics, 2022, 44(1):146-153)
[20] 袁俊,刘国柱,梁宏涛,罗清彩.知识图谱在商业银行风控领域的研究与应用综述.计算机工程与应用 ,2022,58(19):37-52.(Yuan Jun, Liu Guozhu, Liang Hongtao, Luo Qingcai. Summary of Research and Application of Knowledge Graphs in Risk Management Field of Commercial Banks. Computer Engineering and Applications, 2022，58(19):37-52)
[21] 黄珂萍,蒋昌俊.基于本体的城市交通的知识分析和推理.计算机科学,2007，34(3):192-196.(Huang Keping, Jiang Changjun. Analyzing and Reasoning Knowledge of Urban Transportation: Based on Ontology. Computer Science, 2007，34(3):192-196)
[22] 胡浩,刘玉岭,张玉臣等.基于攻击图的网络安全度量研究综述.网络与信息安全学报,2018,4(9):1-16.(Hu Hao, Liu Yuling, Zhang Yuchen, et al. Survey of attack graph based network security metric. Chinese Journal of Network and Information Security, 2018, 4(9): 1-16)
[23] 夏毅,兰明敬,陈晓慧,罗军勇,周刚,何鹏. 可解释的知识图谱推理方法综述.网络与信息安全学报,2022,8(5):1-25.(Xia Yi, Lan Mingjng, Chen Xiaohui, et al. Survey on explainable knowledge graph reasoning methods. Chinese Journal of Network and Information Security, 2022,8(5):1-25)
[24] Maximilian Nickel, Kevin Murphy, Volker Tresp, et al. A Review of Relational Machine Learning for Knowledge Graphs. Proceedings of the IEEE, 2016:11-33.
[25] Wang Q, Mao Z, Wang B, et al. Knowledge Graph Embedding: A Survey of Approaches and Applications.

IEEE Transactions on Knowledge and Data Engineering, 2017, 29(12):2724-2743.
[26] 侯中妮,靳小龙,陈剑赟等.知识图谱可解释推理研究综述[J].软件学报,2022,33(12):4644-4667.(Hou Zhongni, Jin Xiaolong, Chen Jianyun, et al. Survey of Interpretable Reasoning on Knowledge Graph. Journal of Software, 2022, 33(12): 4644-4667)
[27] 马昂,于艳华,杨胜利等.基于强化学习的知识图谱综述.计算机研究与发展,2022,59(08):1694-1722.(Ma Ang, Yu Yanhua, Yang Shengli, et al. Survey of Knowledge Graph Based on Reinforcement Learning. Journal of Computer Research and Development, 2022, 59(8):1694-1722)
[28] Lecuny, Bengioy, Hintong．Deep learning. Nature, 2015, 521(7553):436-444.
[29] Wei Jason, Wang Xuezhi, Dale Schuurmans, et al. Chain-of-Thought Prompting Elicits Reasoning in Large Language Models. Proceedings of the 36th Conference on Neural Information Processing Systems, New Orleans, USA, 2022:1-14.
[30] Bian Ning, Han Xianpei, Sun Le, et al. ChatGPT is a Knowledgeable but Inexperienced Solver: An Investigation of Commonsense Problem in Large Language Models. arXiv: 2303.16421v1.
[31] Aakanksha Chowdhery, Sharan Narang, Jacob Devlin, et al. PaLM: Scaling Language Modeling with Pathways.arXiv:2204.02311v5.
[32] Zhao Wayne Xin, Zhou Kun, Li Junyi. A Survey of Large Language Models. arXiv:2303.18223v9.
[33] 马瑞新,李泽阳,陈志奎等.知识图谱推理研究综述.计算机科学,2022,49(S1):74-85.(Ma Ruixin, Li Zeyang, Chen Zhikui, et al. Review of Reasoning on Knowledge Graph. Computer Science, 2022, 49(S1):74-85)
[34] 张仲伟,曹雷,陈希亮,寇大磊,宋天挺．基于神经网络的知识推理研究综述．计算机工程与应用, 2019, 55(12):8-36.(Zhang Zhongwei, Cao Lei, Chen Xiliang, et al. Survey of Knowledge Reasoning Based on Neural Network. Computer Engineering and Applications, 2019, 55(12):8-36)
[35] 孙水发,李小龙,李伟生等.图神经网络应用于知识图谱推理的研究综述.计算机科学与索,2023,17(01):27-52.(Sun Shuifa, Li Xiaolong, Li Weisheng, et al. Review of Graph Neural Networks Applied to Knowledge Graph Reasoning. Journal of Frontiers of Computer Science and Technology, 2023,17(01):27-52)
[36] 李志飞，赵月，张龑．基于表示学习的知识图谱推理研究综述．计算机科学, 2023, 50(3): 94-113.(Li Zhifei, Zhao Yue, Zhang Yan. Survey of Knowledge Graph Reasoning Based on Representation Learning. Computer Science, 2023, 50(3): 94-113)
[37] 宋浩楠，赵刚，孙若莹．基于深度强化学习的知识推理研究进展综述．计算机工程与应用, 2022, 58(1):12-25.(Song Haonan, Zhao Gang, Sun Ruoying. Developments of Knowledge Reasoning Based on Deep Reinforcement Learning. Computer Engineering and Applications, 2022, 58(1):12-25)
[38] Li Weizhuo, Qi Guilin, Ji Qiu. Hybrid reasoning in knowledge graphs: Combing symbolic reasoning and statistical reasoning. Semantic Web, 2020, 11(1):53-62.
[39] Zhang Wen, Chen Jiaoyan, Li Juan, et al. Knowledge Graph Reasoning with Logics and Embeddings: Survey and Perspective. arXiv:2202.07412v1.
[40] Zhang Jing, Chen Bo, Zhang Lingxi, et al. Neural, symbolic and neural-symbolic reasoning on knowledge graphs. AI Open, 2021, 2:14-35.
[41] Tian Ling, Zhou Xue, Wu Yanping, et al. Knowledge graph and knowledge reasoning: A systematic review. Journal of Electronic Science and Technology, 2022,20(2):100159.
[42] Stefan Schoenmackers, Oren Etzioni, Daniel S Weld, et al. Learning First-Order Horn Clauses from Web Text. Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing, Massachusetts, USA, 2010:1088-1098.
[43] Salmon W, Jeffrey R, Greeno J. Statistical explanation & statistical relevance. 1. Commonwealth of Pennsylvania America: University of Pittsburgh Press, 1971.
[44] Niels Landwehr, Kristian Kersting, Luc De Raedt. Integrating Na¨ıve Bayes and FOIL. Journal of Machine

chinaXiv:202305.00174v1

Learning Research, 2007, 8:481-507.
[45] Niels Landwehr, Andrea Passerini, Luc De Raedt, et al. Fast learning of relational kernels. Machine Learning, 2010, 78:305-342.
[46] Zeng Qiang, Jignesh M Patel, David Page. QuickFOIL: Scalable Inductive Logic Programming. Proceedings of the 2014 International Conference on Very Large Data Bases Endowment, Hangzhou, China, 2014:197-208.
[47] Galárraga Luis, Teflioudi Christina, Hose Katja, et al. AMIE: Association Rule Mining under Incomplete Evidence in Ontological Knowledge Bases. Proceedings of the 2013 International World Wide Web Conference Committee, Rio de Janeiro, Brazil, 2013:413-422.
[48] Galárraga L, Telioudi C, HOSE K, et al. Fast rule mining in ontological knowledge bases with AMIE+. Proceedings of the International Journal on Very Large Data Bases, Hawaii, USA, 2015, 24(6): 707-730.
[49] Lajus J, Galárraga L, Suchanek F. Fast and exact rule mining with amie 3. Proceedings of the European Semantic Web Conference, Hersonissos, Greece, 2020: 36-52.
[50] Wang Zhichun, LI Juanzi. RDF2Rules: Learning Rules from RDF Knowledge Bases by Mining Frequent Predicate Cycles. arXiv:1512.07734.
[51] Meilicke C, Fink M, Wang Yanjie, et al. Fine-grained evaluation of rule-and embedding-based systems for knowledge graph completion. International semantic web conference. Springer, Cham, 2018: 3-20.
[52] Meilicke C, Chekol M W, Ruffinelli D, et al. Anytime Bottom-Up Rule Learning for Knowledge Graph Completion. Proceedings of the International Joint Conferences on Artifical Intelligence, Macao, China, 2019: 3137-3143.
[53] Christian Meilicke, Melisachew Wudage Chekol, Manuel Fink, Heiner Stuckenschmidt. Reinforced Anytime Bottom Up Rule Learning for Knowledge Graph Completion. arXiv:2004.04412.
[54] Ott S, Meilicke C, Samwald M. SAFRAN: An interpretable, rule-based link prediction method outperforming embedding models.arXiv:2109.08002.
[55] Sebastien Ferre. Link Prediction in Knowledge Graphs with Concepts of Nearest Neighbours. Proceedings of the European Semantic Web Conference. Portoroz, Slovenia, 2019:84-100.
[56] Kok S, Domingos P. Learning the structure of Markov logic networks. Proceedings of the 22nd International Conference on Machine Learning, Bonn, Germany, 2005: 441-448.
[57] De-Raedt L, Kimmig A, Toivonen H. A probabilistic prolog and its application in link discovery. Proceedings of the 20th International Joint Conference on Artifical Intelligence, San Francisco, USA, 2007:2462-2467.
[58] Abdus Salam, Rolf Schwitter, Mehmet A Orgun. Probabilistic Rule Learning Systems: A Survey. ACM Computing Surveys, 2021, 54(4):1–16.
[59] 郝娇,惠小静,马硕等.一阶逻辑中公理化真度研究.计算机科学,2021,48(S2):669-671+712.(Hao Jiao, Hui Xiaojing, Ma Shuo, Jin Minghui. Study on Axiomatic Truth Degree in First Order Logic. Computer Science, 2021, 48(S2):669-671+712)
[60] Cussens J. Parameter estimation in stochastic logic program. Machine Learn, 2001, 44(3):245-271.
[61] Wang W, Mazaitis K, Cohen W. Programming with personalized pagerank: a locally groundable first-order probabilistic logic. Proceedings of the 22nd ACM International Conference on Information & Knowledge Management, San Francisco California, USA, 2013: 2129-2138.
[62] Cali A, Gottlob G, Lukasiewicz, T. A general Datalog-based framework for tractable query answering over ontologies. Journal of Web Semantics, , 2012,14:57–83.
[63] Matthias Lanzinger, Stefano Sferrazza, Georg Gottlob, MV-Datalog+-: Effective Rule-based Reasoning with Uncertain Observations. arXiv:2202.01718.

chinaXiv:202305.00174v1

[64] Kimmig A, Bach S H, BroechelerR M, et al. A short introduction
to probabilistic soft logic[C]//Proceedings of the NIPS Workshop on Probabilistic Programming: Foundations and Applications, 2012: 1-4.

[65] Martin Gavalec, Zuzana Němcová, Sergeĭ Sergeev. Tropical linear algebra with the Łukasiewicz T-norm. Fuzzy Sets and Systems, 2014, 276:131-148.

[66] Pujara J, Miao H, Getoor L, et al. Ontology-aware partitioning for knowledge graph identification. Proceedings the Workshop on Automated Knowledge Base Construction, San Francisco California, USA, 2013:19-24.

[67] Bach S H, Broecheler M, Huang B, et al. Hinge-loss markov random fields and probabilistic soft logic. Computer Science, 2017, 18(1): 1-67.

[68] Guo Shu, Wang Quan, Wang Lihong, et al. Jointly Embedding Knowledge Graphs and Logical Rules. Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, Austin Texas, USA, 2016:192–202.

[69] Bordes A, Usunier N, Garciaduran A, et al. Translating embeddings for modeling multi-relational data. Proceedings of the 26th International Conference on Neural Information Processing Systems, New York, USA, 2013:2787-2795.

[70] Wang William, Cohen William. Learning First-Order Logic Embeddings via Matrix Factorization. Proceedings of the 25th International Joint Conference on Artificial Intelligence, New York, USA, 2016:2132–2138.

[71] Guo S, Wang Q, Wang L H, et al. Knowledge graph embedding with iterative guidance from soft rules. Proceedings of the 32th AAAI Conference on Artificial Intelligence, ,2018: 4816-4823.

[72] Trouillon T, Welbil J,Riedel S, et al. Complex embeddings for simple link prediction. Proceedings of the 33rd International Conference on Machine Learning, California, USA, 2016:2071-2080.

[73] Wang P W, Dou D J, Wu F Z, et al. Logic rules powered knowledge graph embedding[J]. arXiv: 1903.03772.

[74] Wang Mengya, Rong Erhu, Zhuo Hankui, et al. Embedding Knowledge Graphs Based on Transitivity and Asymmetry of Rules. Proceedings of the 22nd Pacific-Asia Conference on Knowledge Discovery and Data Mining, Melbourne, Australia, 2018:141–153.

[75] Ding Boyang, Wang Quan, Wang Bin, et al. Improving Knowledge Graph Embedding Using Simple Constraints. Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics, Melbourne, Australia, 2018:110–121.

[76] Niu Guanglin, Zhang Yongfei, Li Bo, et al. Rule-Guided Compositional Representation Learning on Knowledge Graphs. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(03):2950-2958.

[77] Tang Xiaojuan, Zhu Song-Chun, Liang Yitao, et al. RulE: Neural-Symbolic Knowledge Graph Reasoning With Rule Embedding. arXiv:2210.14905v1.

[78] Zhang Wen, Bibek Paudel, Wang Liang, et al. Iteratively Learning Embeddings and Rules for Knowledge Graph Reasoning. Proceedings of the World Wide Web Conference, New York, USA, 2019:2366-2377.

[79] Ho V T, Stepanova D, Gad-Elrab M H, et al. Rule Learning from Knowledge Graphs Guided by Embedding Models. Proceedings of International Semantic Web Conference, Monterey, USA, 2018:72-90.

[80] Omran P G, Wang K, Wang Z. An Embedding-Based Approach to Rule Learning in Knowledge Graphs. IEEE Transactions on Knowledge and Data Engineering. 2021, 33(4):1348-1359.

[81] Cheng Kewei, Yang Ziqing, Zhang Ming, et al. UniKER: A Unified Framework for Combining Embedding and Definite Horn Rule Reasoning for Knowledge Graph Inference. Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing, Punta Cana, Dominican Republic, 2021:9753–9771.

[82] Wang Peifeng, Han Jialong, Li Chenliang, et al. Logic Attention Based Neighborhood Aggregation for

chinaXiv:202305.00174v1

chinaXiv:202305.00174v1

Inductive Knowledge Graph Embedding. Proceedings of the 19th AAAI Conference on Artificial Intelligence, Hawaii, USA, 2019:7152-7159.

[83] Sun Haitian, Bhuwan Dhingra, Manzil Zaheer, et al. Open Domain Question Answering Using Early Fusion of Knowledge Bases and Text. Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing,Brussels, Belgium, 2018:4231–4242.

[84] Hochreiter S, Schmidhuber J. Long Short-Term Memory. Neural Computation, 1997, 9(8):1735-1780.

[85] Wang Hanzhi, Wei Zhewei, Gan Junhao, et al. Personalized PageRank to a Target Node. Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, Virtual Event, CA, USA, 2020:657–667.

[86] He Yongquan, Wang Zihan, Zhang Peng, et al. VN Network: Embedding Newly Emerging Entities with Virtual Neighbors. Proceedings of the International Conference on Information and Knowledge Management, Virtual Event, Ireland, 2020: 505-514.

[87] Zhao Yunxiang, Qi Jianzhong, Liu Qingwei, et al. WGCN: Graph Convolutional Networks with Weighted Structural Features. Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval, Virtual Event, Canada, 2021: 624-633.

[88] Cho Kyunghyun, Bart van Merrienboer, Caglar Gulcehre, et al. Learning Phrase Representations using RNN Encoder–Decoder for Statistical Machine Translation. Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), Doha, Qatar, 2014: 1724–1734.

[89] Qu Meng, Tang Jian. Probabilistic Logic Neural Networks for Reasoning. Proceedings of the 33rd Conference on Neural Information Processing Systems, Vancouver, Canada, 2019: 7712–7722.

[90] Zhang C, Bütepage J, Kjellström H. Advances in Variational Inference. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2019: 41(8):2008-2026.

[91] Veličković P, Casanova A, Liò P. Graph attention networks. Proceedings of the 6th International Conference on Learning Representations, Vancouver CANADA, 2018: 1-12.

[92] Vardhan Harsha, Vivek L, Guo Jia, et al. Probabilistic Logic Graph Attention Networks for Reasoning. Proceedings of the International World Wide Web Conference Committee, Taipei, Taiwan, 2020:669–673.

[93] Zhang Yuyu, Chen Xinshi, Yang Yuan, et al. Efficient Probabilistic Logic Reasoning with Graph Neural Networks. Proceedings of the International Conference on Learning Representations, Kigali, Rwanda, 2023:1-20.

[94] Lei Deren, Jiang Gangrong, Gu Xiaotao, et al. Learning Collaborative Agents with Rule Guidance for Knowledge Graph Reasoning. Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing, Online, 2020:8541–8547.

[95] Xia Yi, Lan Mingjing, Luo Junyong. Iterative rule-guided reasoning over sparse knowledge graphs with deep reinforcement learning. Information Processing and Management, 2022, 59:103040.

[96] Gallaire Herve, Minker Jack, Nicolas Jean-Marie. Logic and Databases: A Deductive Approach. ACM Computing Surveys. 1984, 16(2):153-185.

[97] Tim Rocktäschel, Sebastian Riedel. End-to-End Differentiable Proving. Proceedings of the 31st International Conference on Neural Information Processing Systems, Long Beach, USA, 2017:3791–3803.

[98] Pasquale Minervini, Matko Bosnjak, Tim Rocktaschel, et al. Differentiable Reasoning on Large Knowledge Bases and Natural Language. Proceedings of the AAAI Conference on Artificial Intelligence, New York, USA, 2020:5182-5190.

[99] Pasquale Minervini, Sebastian Riedel, Pontus Stenetorp, et al. Learning reasoning strategies in end-to-end differentiable proving. Proceedings of the 37th International Conference on Machine Learning, Vienna, Austria, 2020:6938-6949.

[100] Cohen W W. Tensorlog: a differentiable deductive database. arXiv: 1605.06523.

[101] Yang Fan, Yang Zhilin, Cohen William W. Differentiable learning of logical rules for knowledge base reasoning. Proceedings of the Advances in Neural Information Processing Systems, LA, USA, 2017:2319-2328.

[102] Wang Po-Wei, Daria Stepanova, Csaba Domokos. Differentiable learning of numerical rules in knowledge graphs. Proceedings of the 8th International Conference on Learning Representations, Virtual, 2020:1-12.

[103] Ali Sadeghian, Mohammadreza Armandpour, Ding Patrick, et al. DRUM: End-To-End Differentiable Rule Mining On Knowledge Graphs. Proceedings of the 33rd International Conference on Neural Information Processing Systems, Vancouver, Canada, 2019:15347-15357.

[104] Yang Yuan, Song Le. Learn to explain efficient via neural logic inductive learning. Proceedings of the 8th International Conference on Learning Representations, Virtual, 2020:1-15.

[105] Ashish Vaswani, Noam Shazeer, Niki Parmar, et al. Attention Is All You Need. Proceedings of the 31st International Conference on Neural Information Processing Systems,Long Beach, USA, 2017:6000-6010.

[106] Dong Honghua, Mao Jiayuan, Lin Tian, et al. Neural Logic Machines. Proceedings of the 7th International Conference on Learning Representations, Virtual, 2019:1-22.

[107] Chen Hanxiong, Li Yunqi, Shi Shaoyun, et al. Graph Collaborative Reasoning. Proceedings of the 33rd International Conference on Neural Information Processing Systems, Vancouver, Canada, 2019:75-84.

[108] Teru Komal K, Denis Etienne G, Hamilton William L. Inductive relation prediction by subgraph reasoning. Proceedings of the 37th International Conference on Machine Learning, Vienna, Austria, 2020:9448-9457.

[109] Mai Sijie. Zheng Shuangjia. Yang Yuedong, et al. Communicative Message Passing for Inductive Relation Reasoning. Proceedings of the AAAI Conference on Artificial Intelligence, Vancouver, Canada, 2021:4294-4302.

[110] Jiajun Chen, Huarui He, Feng Wu, et al. Topology-Aware Correlations Between Relations for Inductive Link Prediction in Knowledge Graphs. Proceedings of the 35th AAAI Conference on Artificial Intelligence, Vancouver, Canada, 2021:6271-6278.

[111] Lin Qika, Liu Jun, Xu Fangzhi, et al. ConGLR: Incorporating Context Graph with Logical Reasoning for Inductive Relation Prediction. Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval, Madrid, Spain, 2022:893–903.

[112] Xu Xiaohan, Zhang Peng, He Yongquan, et al. Subgraph Neighboring Relations Infomax for Inductive Link Prediction on Knowledge Graphs. Proceedings of the 31st International Joint Conference on Artificial Intelligence Main Track, Shenzhen, China, 2022:2341-2347.

[113] Zha Hanwen, Chen Zhiyu, Yan Xifeng, et al. Inductive Relation Prediction by BERT. Proceedings of the AAAI Conference on Artificial Intelligence, Virtual, 2022:5923-5931.

[114] Pan Yudai, Liu Jun, Zhang Lingling, et al. Learning First-Order Rules with Relational Path Contrast for Inductive Relation Reasoning. arXiv:2110.08810.

[115] Yan Zuoyu, Ma Tengfei, Gao Liangcai, et al. Cycle Representation Learning for Inductive Relation Prediction. Proceedings of the International Conference on Learning Representations, Virtual, 2022:1-16.

[116] Qu Meng, Chen Junkun, Louis-Pascal Xhonneux, et al. RNNLogic: learning logic rules for reasoning on knowledge graphs. Proceedings of the 2021 International Conference on Learning Representations, Virtual, 2021.

[117] Huang Qian, Ren Hongyu, Jure Leskovec. Few-shot Relational Reasoning via Connection Subgraph Pretraining. Proceedings of the 36th Conference on Neural Information Processing Systems, New Orleans, LA, USA, 2022:6397-6409.

[118] Du Zhengxiao, Zhou Chang, Yao Jiangchao, et al. CogKR: Cognitive Graph for Multi-Hop Knowledge Reasoning. IEEE Transactions on Knowledge and Data Engineering. 2021, 35(2):1283-1295.

[119] Liu Y, Ma Y, Hildebrandt M, et al. TLogic: Temporal Logical Rules for Explainable Link Forecasting on Temporal Knowledge Graphs. Proceedings of the AAAI Conference on Artificial Intelligence, Virtual,

chinaXiv:202305.00174v1

2022:4120-4127.
[120] Bang Yejin, Samuel Cahyawijaya, Nayeon Lee, et al. A Multitask, Multilingual, Multimodal Evaluation of ChatGPT on Reasoning, Hallucination, and Interactivity. arXiv:2302.04023.
[121] Toutanova K,Chen D. Observed versus latent features for knowledge base and text inference. Proceedings of the 3rd Workshop on Continuous Vector Space Models and their Compositionality, Beijing, China, 2015:57-66.
[122] Dettmers T,Minerbini P,Stenetorp P,et al. Convolutional Knowledge Graph Embeddings. Proceedings of the 32nd AAAI Conference on Artificial Intelligence, New Orleans, USA, 2018:1811-1818.
[123] Xiong W, Hoang T, Wang W Y. DeepPath:A Reinforcement Learning Method for Knowledge Graph Reasoning. Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, Copenhagen, Denmark, 2017-564-573.
[124] Williams R J. Simple statistical gradient-following algorithms for connectionist reinforcement learning. Machine Learning. 1992, 8(3–4):229–256.
[125] Nickel Maximilian, Tresp Volker, Kriegel Hans-Peter. A Three-Way Model for Collective Learning on Multi-Relational Data. Proceedings of the the 28th International Conference on International Conference on Machine Learning, Bellevue Washington, USA, 2011:809–81.
[126] Schlichtkrull M, Kipf T N, Bloem P, et al. Modeling Relational Data with Graph Convolutional Networks. Proceedings of the 15th European Semantic Web Conference, Heraklion, Greece, 2018:10843.

(通讯作者：成清　E-mail:zzfzad2021@163.com)

作者贡献声明：

曾泽凡：文献调研分析、论文起草、最终版本修订修订；
司悦航：文献调研；
刘忠：论文写作指导；
成清：研究命题的提出、设计。

chinaXiv:202305.00174v1